PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : G06F 15/38	A1	(11) International Publication Number: WO 93/25974
		(43) International Publication Date: 23 December 1993 (23.12.93)

(21) International Application Number: PCT/US93/05561

(22) International Filing Date: 10 June 1993 (10.06.93)

(30) Priority data:
07/897,362 11 June 1992 (11.06.92) US

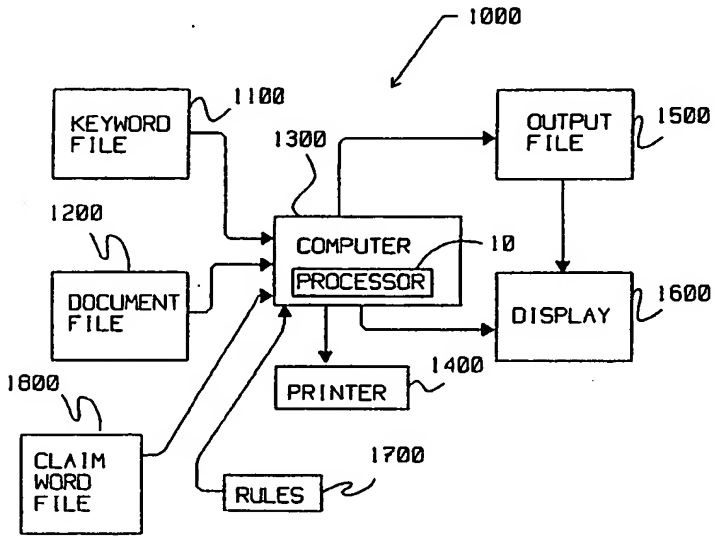
(71)(72) Applicant and Inventor: MOFFA, Emil [US/US]; 1903 County Road D, Arden Hills, MN 55112 (US).

(74) Agent: LEONE, George, A., Sr.; Leone & Moffa, 1300 Godward Street, N.E., Suite 6500, Minneapolis, MN 55413 (US).

(81) Designated States: AU, BB, BG, BR, CA, CZ, FI, HU, JP, KP, KR, LK, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SK, UA, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: AN AUTOMATED METHOD FOR CHECKING PATENT APPLICATIONS



(57) Abstract

An automated method for checking patent applications or documents (1000) includes the steps of identifying subgroups of serial characters which relate to each other from within serial group of characters (200-350) and checking the subgroups for consistency of relation with regard to drafting rules (1700). Reference characters (600) are checked for consistency. A patent application is checked for all required parts (700-760). Inconsistency among elements for a selected character (600) is checked. Claim section (740), the number of claims (811) with proper dependency (920-948) are also checked. A recited hierarchical relationship is used to build a claim structure for checking antecedent basis for a family of claims (984). Elements needing antecedent basis are isolated and checked against recited potential antecedents appearing in the proper order (824-832). Claim elements lacking antecedents are reported to the user (830-834).

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**MR**	Mauritania
AU	Australia	**GA**	Gabon	**MW**	Malawi
BB	Barbados	**GB**	United Kingdom	**NL**	Netherlands
BE	Belgium	**GN**	Guinea	**NO**	Norway
BF	Burkina Faso	**GR**	Greece	**NZ**	New Zealand
BG	Bulgaria	**HU**	Hungary	**PL**	Poland
BJ	Benin	**IE**	Ireland	**PT**	Portugal
BR	Brazil	**IT**	Italy	**RO**	Romania
CA	Canada	**JP**	Japan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SK**	Slovak Republic
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SN**	Senegal
CM	Cameroon	**LI**	Liechtenstein	**SU**	Soviet Union
CS	Czechoslovakia	**LK**	Sri Lanka	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**MC**	Monaco	**UA**	Ukraine
DK	Denmark	**MG**	Madagascar	**US**	United States of America
ES	Spain	**ML**	Mali	**VN**	Viet Nam
FI	Finland	**MN**	Mongolia		

An automated method for checking patent applications

FIELD OF THE INVENTION

The present invention relates to the field of composition, and in particular, to an automated method and apparatus for checking compliance of a serial group of words such as a drafted document against a set of word grouping rules such as drafting rules. More particularly, the invention relates to an automated method and apparatus for checking a patent application against the rules for drafting a patent application.

BACKGROUND OF THE INVENTION

Legal documents, as, for example, patent applications have long been drafted and corrected by attorneys through a process of drafting, reading and revising. Patent applications in particular are drafted in accordance with specific drafting rules following the patent statutes and the Manual of Patent Examining Procedure (MPEP). Such drafting rules include rules for reference characters for elements wherein like elements have like reference characters, rules requiring antecedent basis for claim elements and other drafting rules.

Prior to the automated method provided by the instant invention, the only way to check documents against these drafting rules has been by reading and revising the patent with the reader keeping the rules and the patent under draft in mind while reviewing the patent and indicating changes.

It is, therefore, the motive of the invention to provide an automated method of checking a patent application. **This application was checked with the method of the invention.**

SUMMARY OF THE INVENTION

An automated method for checking a document comprised of a serial group of characters is provided. The method comprises the steps of identifying subgroups of serial characters which relate to each other from within the serial group of characters, and checking the subgroups for consistency of relation with regard to drafting rules.

In one aspect of the invention, the method includes the steps of allocating a unique location for each occurrence of each of a plurality of reference characters and then selecting each reference character in turn and finding at least a first preceding word for each occurrence of each selected reference character. Then the at least first preceding word is identified as an element of the selected reference character. The steps of selecting reference characters and identifying elements are repeated until all occurrences of all reference characters have been found. Then each reference character is selected in turn and all associated identified elements are counted for each selected reference character. All identified elements are compared against each other for each selected reference character. If an inconsistency is found among the elements for a selected reference character an error message is generated.

In another aspect of the invention the method includes the steps of finding parts of the patent application, including the background, field of art, summary, abstract, description of the drawings, description of the preferred embodiment and claims. If a claim section is found the method counts the number of claims and analyses them for proper dependency. The method also uses the recited

hierarchical relationships to build a claim structure that is used to check the claims for antecedent basis. Antecedent basis is checked for within each claim, and within claims depending from other claims. Elements needing antecedent basis are isolated then checked against recited potential antecedents appearing in the proper order. Claim elements lacking antecedents are reported to the user.

The user can control the level of verbosity and amount of error reporting for each error type. The output of the analysis is available to the user in ASCII format, in an interactive query format, and is available for further processing in binary or ASCII.

Other objects, features and advantages of the invention will be apparent through the Description of the Preferred Embodiment, Claims and Drawings herein wherein like character refer to like elements.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a simplified block diagram generally showing the main components of the automated method and apparatus of the invention for checking a document against a preselected set of drafting regulations.

Figures 2A and 2B are intended to be read together as a process flow diagram of one embodiment of the automated method of the invention for checking a document against a set of drafting rules.

Figures 3A and 3B are intended to be read together as a process flow diagram of one embodiment of the invention for building word strings from a document file.

Figure 4 is a process flow diagram illustrating one example as provided by the automated method and apparatus of the invention for checking reference characters against words in a document.

Figure 5 shows a word structure as employed in

one embodiment of the invention.

Figure 6 shows a reference character structure as employed in one embodiment of the invention.

Figure 7 shows a process flow diagram of one method of automatically reading a patent document as employed in one embodiment of the invention.

Figure 8 is a process flow diagram of an automated method of reading the claims section of a patent document for identifying claimed elements.

Figure 9 shows a process flow diagram for finding all the claims of a patent specification.

Figure 10 shows a process flow diagram of the method of the invention to determine claim dependencies.

Figure 11 shows the method of the invention used to check the consistency of the references to drawings in the brief description of the drawings section with the remainder of the patent specification.

Figure 12 shows the method of the invention used to read in claim words.

Figure 13 shows the method of the invention used to assign claim word status.

Figure 14 shows the method of the invention used to read in trigger words.

Figure 15 shows the method of the invention to find claim elements.

DESCRIPTION OF THE PREFERRED EMBODIMENT

For the purposes of explaining the invention herein, the following terms shall have the following meanings herein unless otherwise indicated by the context in which they are used.

TERMS	DEFINITION
ANTECEDENT	The first recitation of a claim element within a claim or series of claims. An antecedent is comprised of a word string which has a first and last word.
WORD	Any string of characters in a document which represents a word in a selected language.
WORD STRING	A set of consecutive words.
CHARACTER STRING	A set of consecutive characters.
ELEMENT	A word or word string referring to a particular illustrated feature in a patent document.
REFERENCE CHARACTER	A designator which identifies an element in a patent document.
CLAIM ELEMENT	A word string within a claim which represents a claimed feature, function, step, part, member or other basic structure.
CLAIM WORDS	Words or symbols predefined as delimiters when used in a patent claim. See examples listed hereinbelow.

Referring now to Figure 1 which shows a block diagram conceptually showing the main components of a system for checking a document against a preselected set of drafting regulations. The system 1000 includes

a keyword file 1100, a document file 1200, and rules file 1700. The keyword file 1100, document file 1200 and rules file 1700 are accessed by a computer 1300 which includes a processor 10 wherein the processor 10 operates on the rules, documents and key words to check the document against the rules from the rules file 1700. The operations of processor 10 are explained in detail hereinbelow.

When the document file 1200 has been processed by processor 10 an output file 1500 and a claim word file 1800 may be created by the computer. The computer 1300 may also display information on a display device 1600 or print information on another output device such as printer 1400. The contents of the various files are described in more detail hereinbelow. The computer 1300 may, advantageously be any of a wide variety of commercially available computers such as a personal computer based upon an INTEL 386 or 486 microprocessor chip. The display device 1600 and printer 1400 may be any commercially available printer or display suitable for use with a personal computer. The rules file 1700 is optional and may be supplanted by a computer program within the processor 10.

Generally, in a law office setting patent specifications are available as word processing documents prepared by word processing programs. These programs have the ability to generate an ASCII file of the patent specification or a binary file of the patent specification. The patent specification can be analyzed with the method and apparatus of the invention in either form. Mistakes found in the patent specification may be directly corrected by overwriting or inserting the automatically detected errors. This correction is done by modifying the document file and replacing certain elements of the

document file. For example, if a recited element has an associated drawing number and it is used in the specification without that drawing number, the drawing number can be supplied. The supplied drawing number is written directly to the document file 1200. Once modified by the method and apparatus of the invention the document file can then be reprocessed to verify that the modification was correct.

The method and apparatus of the invention used to automate the checking of a serial group of words against word grouping rules reads the serial group of words and constructs a logical structure that enables the analysis of consistency assertions. These consistency assertions are within the scope of the serial groups of words themselves, these words are used to generate a list of elements that have relationships to each other that are consistently used throughout the serial group of words. Thus the serial group of words themselves provide the context in which the document is analyzed. The keyword file 1100 and claim word file 1800 provide a way of establishing the parameters of the document that override the self consistency requirements of the serial group of words. Certain words have specific meanings in the field of patent law. These meanings override the use of these words within the document's self consistency. Those skilled in the art will recognize that other fields such as drafting of contracts, or licenses follow similar self consistency restrictions in that terms and objects referenced in those documents need to be self consistent in their use.

Figures 2A and 2B may be read pieced together to show one illustrative example of the invention as used in the specific application of automatically checking a patent application or patent document against

drafting rules and other legal requirements. Those skilled in the art will recognize that, while the invention will be explained in terms of automatically checking a patent document, it is not so limited and other applications of the invention for checking other types of documents will come within the spirit and scope of the claims herein.

Referring now in particular to Figure 2A, the first process step after initiating the process of the invention is to load keywords 20 from the keyword file 1100. Keywords are designated in the keyword file 1100 and are defined as common words which occur frequently in proximity to numbers. This is done to avoid confusion with reference characters used to designate steps or elements of a process or apparatus. A partial table of some examples of keywords is shown below in Table 1. Each keyword is followed by a pointing character "n" or "b".

TABLE 1

KEYWORDS
a n
a b
the n
the b
about n
was n
at n
of n
of b
with b
with n
are n
are b
then n
then b

```
than n
than b
every n
every b
least n
figure b
figures b
sigma b
lb n
lbs n
mps n
ms n
meters n
ft n
volt n
volts n
duty n
cycle n
degree n
degrees n
bit n
bits n
hz n
khz n
vdc n
vac n
arcsec n
arcseconds n
to n
and b
from b
. b
claim b
numeral b
january b
```

january n
february b
february n
march b
march n
april b
april n
may n
may b
june n
june b
july n
july b
august n
august b
september n
september b
october n
october b
november n
november b
december n
december b
members b
all n
samples b
table n
-END TABLE 1-

In the keyword file 1100 listed above the character "n" following a keyword stands for "next" and indicates that a number may follow the selected keyword. The character "b" following a keyword stands for "back" and indicates that a number may precede the selected keyword. By recognizing keywords and the "n" and "b" pointing characters the process of the

invention recognizes numbers such as dates and measurements and can differentiate such keyword associated numbers from reference characters.

In an alternate embodiment of the invention the keyword file can be encoded in a binary fashion with a three state delimiter either indicating that the number follows the word, the number precedes the word or the number can follow or precede the word. Those skilled in the art having the benefit of this disclosure will also recognize that different keyword files may be used depending on the type of art being disclosed in the patent specification. The user may provide keywords that can be used for a specific patent application to adjust the sensitivity of the automatic method and apparatus of the invention. The apparatus of the invention can learn keywords as more and more patent applications are examined by receiving feedback from the user of the invention as to whether or not the subject word causing an indicated error is indeed a keyword.

If no keyword file 1100 is used, a message to that effect may optionally be printed or displayed by the apparatus of the invention at process step 30. A default set of keywords may be used if desired. A keyword file 1100 is optional at the discretion of the user.

In one embodiment of the invention, the next process step 40 reads the parts of the patent document. Process step 40 is explained in detail below with reference to Figure 7. After the patent has been read the process moves to the next process step 50 wherein a predetermined number of characters is printed from the document which was read in step 40. In the case of a patent document, for example, a paragraph or two of descriptive text may be printed to

aid the reader in understanding a report which is generated by the process and apparatus of the invention. After reading and, optionally printing a portion of the patent document in step 50, the process moves to process step 60 wherein all words in the document are analyzed as to their intrinsic type.

Process step 60 may also include an optional tokenization of the words. Tokenization is an optional step which may be employed to achieve compiling speed and a degree of data compression using a hashing scheme. As part of the tokenization step, for example, the document words may each be assigned a hash code in accordance with well known techniques for optimizing computer programs.

In the next process step 70 the apparatus and method of the invention reads the document for reference characters and puts each occurrence of a reference character in a list. The method and apparatus of the invention proceeds to process step 80 wherein each word is checked against each reference character found in step 70. Next at process step 90 elements associated with each reference character are compared against each other for each reference character in turn. Elements associated with a selected reference character which do not match other elements associated with the same selected reference character are flagged as errors for each reference character in step 100.

Optionally, the method of the invention may check for words matching element word strings which have no associated reference character at step 110. Elements missing reference characters may be reported as errors. The process then passes to step 120 wherein, in another optional subprocess, the process of the invention uses information from process step 40 to

isolate the claims portion of the patent document and counts the number of claims. Once the number of claims have been counted the process moves to step 130 wherein the number of independent claims is identified and counted and then on to step 140 wherein the number of dependant claims are identified and counted. As part of process step 140 the root independent claims and any dependent claims from which other dependant claims depend in a chain-like fashion are also identified. A claims report may advantageously be printed at process step 150. At process step 160 the claims portion is checked for antecedents. The time and date are stamped on any output at process step 170.

Once the errors in the document have been detected the user is notified. In one embodiment of the invention only the first error of a mislabeled element is reported to the user. In an alternate embodiment of the invention all errors in all instances of usage are reported to the user. The user can select the level of verbosity of the resulting output. As indicated earlier the user can also request that the system automatically correct the specification by, for example, supplying the correct reference character in place of the incorrect reference character.

In one embodiment of the invention a predetermined list of reference characters and elements used in the document is generated. Any inconsistency from the generated list is reported to the user. The generated list of reference characters and elements may be keyed to the patent drawings to reference like elements with like reference characters.

Those skilled in the art will recognize that the

data structures used for the preferred embodiment exploit the features of the "C" language. Other languages may be used to implement the method and apparatus of the invention including object oriented languages such as "C++".

Referring now to Figure 5, a word structure 500 as employed in one embodiment of the invention is shown. The word structure 500 includes pointers to addresses of selected words in the documents. Those skilled in the art will understand that the invention is not limited to such a structure but that, in some applications, data may be substituted for address pointers. In the example shown in Figure 5, the word structure 500 comprises the fields *BACK 510, *NEXT 520, *STRING 530, *TYPE 540 and LOCATION 550. The example shown is taken from an embodiment of the invention implemented in the C programming language. Therefore, the parameters preceded by an asterisk such as *BACK are pointers for that type, where type here is being used according to the definition of a type in the C programming language. For each word in the document therefore, a word structure is created including *BACK which points to the word preceding, as in looking back, from the current word or reference character. The parameter *NEXT points to the next word or other character string following the current word being processed. The parameter *STRING represents a pointer to a storage location for a document element. A character string is usually a word in the document but may also be a numeral, punctuation, a keyword, or a reference character. Once the processor 10 has processed a word string it creates a pointer called *TYPE. *TYPE is a pointer to corresponding data designating that the string pointed to by *STRING is a document element type, as for

example, a keyword, number, reference character or word. The processor 10 builds each string one character at a time, as described herein checking each character to determine whether it is a letter, punctuation mark, null character, space, carriage return or letter. Those skilled in the art will appreciate that other compiling methods may be used such as a stack oriented approach.

Referring now to Figure 6, a reference character structure 600 as employed in one embodiment of the invention is shown. The reference character structure 600 is comprised of the following elements: *NEXT 610 is a pointer to the next element of the reference character structure thus creating a linked list of reference numbers; VALUE 620 is an integer representing the numerical value of the reference character; and *WORD 630 is a pointer to the location of the reference character within the document. Since the document is represented as a linked list of words, this pointer also points to the word representation of the character. Thus the word "ten" would have a value of 10. In an alternate embodiment of the invention *WORD may point to the word before or after the number.

As in the case of words, reference characters are built on a character by character basis by examining a character as it is processed, herein conveniently referred to as "C", as detailed herein.

Referring now to Figures 3A and 3B a flow diagram for a process and apparatus of accessing words from a document is shown. This relates to step 60 in the flow diagram of Figures 2A and 2B. Initially at step 200, the process checks for the presence of the document file to be processed. The process exits at step 210 if the file is not available, or if the file

is a null file. If the file can be accessed the process designates the last accessed character, "LAST" as NULL or equivalently zero. At step 215 a determination is made as to whether or not the last accessed character, LAST, is equal to zero. If "LAST" is not zero the process branches to step 220. Under the convention used in this example, if LAST = 0, then the next character is fetched from the file as character "C" at step 230. If LAST is not zero, then the next character to be processed, "C", is set equal to LAST at step 220. The last accessed character from the previous pass through the process starts the next character string to be processed if the last accessed character has not been consumed. LAST is reset to zero for the next pass through the process if the last accessed character has been consumed. At step 235 the location of "C" is noted with respect to its position in the document.

The process then proceeds to step 250 where an array called "STRING" is set to null in its first indexed position STRING[0] in this example. The rest of process 250 as explained below then proceeds to fill the STRING[index] array character by character by operating on the currently selected character, "C". The process flows to step 260 wherein the word being built is read for punctuation. If the current character "C" at 270 is determined to be punctuation then the STRING[index] is allocated as punctuation using the designator "P" and the process routes to step 350 where the special case of "C" being a period is considered wherein a period is differentiated from a decimal point. If the current character "C" is not punctuation then the program at process step 280 changes the character "C" to lower case if it is an upper case letter. This step can be omitted if case

sensitivity is desired. The process then flows to step 290 where the last access character is set equal to the current character in order to preserve the last character. Step 290 may be avoided if other well-known storage means are employed to preserve the last character. At step 300 the string array is terminated when a NULL character is encountered. The string is returned as a word and not a reference character if the first element of the string is not a number character between 0-9. Until a null character is fetched, the process continues to cycle through steps 215 through 300 reading the document.

At step 320 the current STRING[index] array is checked character by character to verify whether it is a reference character as opposed to, for example a number indicating a quantity or other value. The process then flows to step 330 wherein if the string is not a reference character the string is allocated as a word. If, however, the first character of the string is a number character between 0-9, the string is allocated as a reference character at step 340.

Now referring to Figure 4, a detailed flow diagram is shown for process steps 80 and 90 wherein elements are checked for proper association with their reference character. As indicated above, at step 80 the process checks each word in the document against each element of the document that have an associated reference character. In one alternate embodiment of the invention the document words are not stored in random access memory but are maintained as an analytical structure on a mass storage unit. This enables the analysis of documents that may not fit into available random access memory. On systems with virtual memory support this is less of a problem.

Prior to checking, the document is spell-checked

to avoid spelling errors. In an alternate embodiment of the invention the comparison of document elements would also consider aliases, abbreviations and similar words. Thus in one field of art "process step 122" may be considered equivalent to "process block 122" obviating the need to alert the user to the harmlessly inconsistent usage. Likewise "jet plane" could be shortened to "jet" without causing an error, etc.

The process then flows to step 360 wherein a stack of occurrences of each reference character is built. Next at step 370 the method and apparatus of the invention checks the character string preceding the numeral, "BACK", against the keyword table. If "BACK" is the equivalent of a keyword the process branches to step 375 wherein the "BACK" word is not counted as an element. If the "BACK" character string is not a keyword the process flows to step 380 wherein the "BACK" word is counted as an element. At process step 390 the number of elements is compared against the Null set. If the number of elements is Null the reference character will be reported to the calling process as having no elements and the next reference character will be obtained at process step 400 returning the process flow to step 80. If the number of elements is not the Null set, this indicates that there are other elements in the set of elements defined by a particular reference character. The process then reads through all words in the document looking for all elements at step 410. The process then flows to step 420 wherein it creates a count of elements for each reference character. The process then flows to process step 430 wherein an array of elements is dynamically allocated for each reference character. At process step 440 the words in the document corresponding to each elements' pointers are

compared. For each reference character as determined at process step 320, the set of elements are compared against each other. Each occurrence of non-identical elements for a selected identical numeral results in a non-comparing element. All "non-comparing" elements are reported as errors for the selected reference character.

As stated above, the sensitivity of comparison may be modified by taking into account the use of an alias, an abbreviation, or equivalent words predefined by the user.

Referring now to Figure 7 a more detailed flow diagram of the method of reading the parts of a patent document is shown. The steps include finding a background section 700, finding a summary 710, finding a brief description of the drawings 720, finding a description of the preferred embodiment 730, and finding a claims section 740. If any of the sections are missing a report is generated at step 750 announcing that the particular section has not been found. If all of the sections are found and, optionally, if at least one of the pertinent sections is found selected parts of the patent document are printed at step 760. The processor 10 finds each section by finding word strings which match predefined section headers. These section headers may include for example, "BACKGROUND OF THE INVENTION", "SUMMARY OF THE INVENTION", "BRIEF DESCRIPTION OF THE DRAWINGS", "DESCRIPTION OF THE PREFERRED EMBODIMENT", and "I CLAIM". If word strings within the patent document are not found to match the predefined headers, a report is generated stating that the particular section has not been found. For each section found, a pointer is reserved to locate the section if needed by the processor 10 for further

processing. For example, the pointer at which the claims start is reserved for use later on in locating the claims section when it is necessary to check the claims.

Referring now to Figure 8, a flow diagram of one example of a method for identifying claimed elements in a patent is disclosed. At process step 800 the method starts. In process step 802 the processor searches for the claims section from the other parts of the patent previously identified as described hereinabove with reference to Figure 7. Once the claims section has been located, as by a pointer or other indexing means well known in the art, the process proceeds to step 804. Optionally, the location of the first claim may be obtained absent the inclusion of the words "I claim" or the like by searching for claim like language, the method of which is disclosed below.

At process step 804 the claim section is either found or it is not found. If the claims section is not found the process ends at step 805. If the claims section is found the process advances to step 811 to find the next claim. If a next claim is found the process advances to step 810. If a next claim is not found then the process ends at step 805. In step 810 the antecedent method of the invention finds the next claim element. If there are no more claim elements the process returns to step 811. If there is a claim element the process advances to step 820 to check if the claim element is a claim antecedent or needs a claim antecedent. If the claim element is a potential antecedent the process flows to process step 822. If the claim element needs an antecedent basis the process flows to process step 824. In process step 822 the antecedent is checked as to whether it is

already in the antecedent list for this claim or parent claims. If the antecedent is already in the list the process advances to process step 830 to report this error. If the antecedent is not in the list the process advances to process step 826 to add this antecedent to the list of antecedents for this claim.

In process step 824 the claim element is compared to the antecedents already assembled for that claim. If an antecedent is found in the list of antecedents for that claim the process returns to step 810 to check the next claim element. If an antecedent is not found in the list of antecedents the parent claim antecedents are check hierarchically in step 828. If in step 832 an antecedent is found in the list of antecedents for that claim's parents the process returns to step 810 to check the next claim element. If an antecedent is not found in the list of antecedents the error is reported in process step 834 and the process returns to step 810 to check the next claim element.

In one embodiment of the invention a linked list is used to store the antecedent claim structure to permit dynamic memory allocation and deallocation of the antecedent claim structure.

Now referring to Figure 9 which shows a method of finding all the claims in a patent document. The method of Figure 9 is generally associated with process block 811 in Figure 8. The method is an alternate method which could be used to first find all the claims then check their antecedents. The method of Figure 8 finds each claim and after the claim is found the antecedents are checked immediately. The method of finding all claims 811A starts in process block 902. The process flows to block 904 to load the

current claim number which, in one embodiment of the invention is claim 1. The process flows to process step 906 to start scanning at the beginning of the claims section which was found at process block 802 in Figure 8. The claims section was found with the method of Figure 7. The process of Figure 9 then flows to process block 908 to load the next word of the claims section. If the next word of the claims section is null the process flows to process block 916 to note where the end of the claims were and the highest claim number. The process then flows to step 805A. If in process step 908 the loading of the next word of the claims section is not null the process flows to block 910. In process step 910 the next word is checked against the current claim number. This is a check where the words are interpreted to be numbers. If the word has an associated number significance and the numbers are identical to the current claim number in this example "1", the process flows to process block 912. At process block 912 the method changes the current claim number to "2". If, in process check step 910, the words are not the current claim number, or the word is not similar to the next word, the process flows back to step 908 to load the next word of the claim section. In process step 914 the claim number is associated with its location in the patent document. This is to allow the processing of antecedents in a hierarchical fashion. By noting an association between the start of a claim and the location in the patent document antecedents can be checked.

Referring now to Figure 10 which shows a method of determining the dependencies of a claim structure in a patent. The process starts at step 920 and flows to step 922 where the current claim is set to be the

first claim to be checked. In one embodiment of the invention the first claim is claim 1. The process flows to step 924 to check whether or not the claim references a claim within it. In one alternative embodiment of the invention the process reads each word in the claim, the words in each claim being defined as bounded between the current claim and the start of the next claim. These positions are determined by the method of Figure 9. If the claim does not reference a claim within it as indicated by either a specific reference to a single claim or a specific reference to a number of claims indicated by the word "claim" or "claims" within the claim the process flows to step 926 to indicate an independent claim. If there is a reference to a claim then the claim number is first determined as the numerical equivalent of the word after the word "claim" or "claims". The process flows to step 934 to determine whether or not the referenced claim that it depends on is indeed itself. If it is then an error is reported in step 936 to indicate that the claim depends on itself. If the claim is not dependent on itself the process flows to step 938 to check whether or not the claim depends on a non-existent claim. If the claim does depend on a non-existent claim then the error is reported in 940. The process then flows to step 942 to determine whether the claim depends on a higher claim number. If it does the process flows to step 934 to report this error. The process flows to step 946 to indicate that the claim is a dependent claim. The process flows then to step 928 to make the next claim the current claim. The process flows to 930 to check whether this claim is the last claim. If it is not the last claim the process flows to step 924 to check whether or not the current claim makes a

reference to a claim as described above. If it is the last claim the process flows to 932 to report the dependency structure and, in one alternate embodiment of the invention, create a claim dependency structure that associates every claim with the claim or claims that it depends on. The process then ends at step 948.

Now referring to Figure 11 which shows the method and apparatus of the invention used to check the consistency of the drawings within the patent specification. The method of the invention begins at process step 1102 by starting the drawing checking process. In process block 1104 a brief description of the drawing section is located. First, the process attempts to locate the brief description of the drawings specifically by title in section 1106 following the method of the invention described above with reference to Figure 7. If the title is found the process flows to process block 1112 where the specification is read. If the brief description of the drawings section is not found by title then the process flows to process block 1108 to determine the drawing section by the descriptive words of the patent. If by descriptive words the brief description of the drawing section is not found the process ends at process step 1110 in an error report stating, for example, that the brief description of the drawing section has not been found. If in either process case 1108 or 1106 the brief description of the drawing section is found the process flows to process blocks 1112 and 1114.

In process step 1114 the method of the invention generates a figure usage structure for the figures described in the brief description of the drawings. The process checks whether the figures are numbered

sequentially and whether they have an associated description. After generating a figure structure for each figure and storing the structure in a linked list of figures for the brief description of the drawing the process flows to step 1118 to check whether the figures are consistent with the figures from the specification. In process step 1112 the specification is read for the use of figures following the methods described above. For each figure encountered a figure structure is generated in process step 1116. The usage of each figure is put into a link list. The process flows to step 1118 to check the consistency between the figure structure generated while reading the specification excluding the brief description of the drawings with the structure created reading the brief description of the drawings. If each drawing figure has a consistent description and there is a one to one correspondence with each element of both sets then the process flows to 1122 where no errors are reported and optionally the process reports that the figures are consistent. If the figures are not consistent the process flows to section 1120 to report that the figures are inconsistent.

Now referring to Figure 12 which shows the method of the invention loading in claim words. The method defines a list of words known as claim words. The process starts at step 952 and loads the claim words into an array of claim words at step 950. The process ends at step 954. The claim words are words found in a claim that serve to delimit claim elements.

The following words and symbols are examples of claim words as described hereinabove.

CLAIM WORDS

comprising

containing

consisting
,
.
;
:
(
comprises
includes
having
has
for
is
wherein
further
are
adapted
which
and
such
in
along
around
below
above
over
under
beneath
after
before
means
consists

Claim words are used with reference to Figure 15 when reading in claim elements to check for antecedent basis.

Now referring to Figure 13 which shows an

alternate embodiment of the invention used to check whether or not a word in a claim is a claim word. The process of Figure 13 starts at step 956 and flows to step 958 where the claim being checked is determined. For instance, in one example, claim 1 will be checked. The process flows to step 960 where each word in claim 1 is scanned and assigned a claim word status. In this alternate embodiment of the invention the claim word status is used in subsequent processes of the invention to determine whether or not the word in question is indeed a claim word. Claim words are not considered to be part of claim elements. The process advances to step 962 where each word in the claim now has a structure associated with it to indicate whether it could possibly be a claim element or is a claim word.

Now referring to Figure 14 which shows a method of reading trigger words. The process starts at 964 where the trigger words are read in. The process flows to 966 where an array is read of trigger words which serve to trigger the start of a claim element. Some trigger words that are commonly used in claims are "a", "an", "means", "means for", "at", "said" and "the". The process ends at step 968 where the trigger words create an array used in subsequent processes of the invention.

Now referring to Figure 15 which shows the method of the invention used to find claim elements within a claim. The process starts at step 970 where a particular claim, for instance, claim 1 will be examined for antecedent basis. The process 810A of Figure 15 is analogous to step 810 in Figure 8. The process first loads the current word to be the first word in the current claim. The process then flows to step 974 where the word is determined to be a trigger

word. If the word is not a trigger word the process flows to step 986 to check the next word in the claim. The process flows to 988 where if the next word is determined to be null the process ends at step 990. If the next word is not null the process returns to step 974 to check whether the next word is a trigger word. If the word being checked is a trigger word the process flows to 976 to find the next claim word. The process flows to step 978 to find whether or not a claim word is found. If it is not found the process advances to step 980 and indicates an error. If a claim word is found or the end of the claim is found the process flows to step 982 to allocate a new element which will include all words from the word immediately after the trigger word through and including the word immediately before the claim word. Thus the claimed element will be defined between the boundaries of the trigger word and the claim word. The process steps to 982 where the element is allocated and then the process flows to 984 for antecedent basis checking or antecedent basis creation according to the methods of Figure 8. The process then flows to 986 to check the next word in the claim and flows to step 988 to see if it is null. If it is null the process ends at step 990, if it is not, the process goes on to the next word.

In an alternate embodiment of the invention each trigger word may create alternate structures that form a number of additional antecedents beyond that specifically recited in the patent. For example, the phrase "a means for" will generate a number of alternate antecedents. For example, a claim recitation of "computing means for dividing by two" generates antecedents including "computing means" and "means for dividing by two", "computing means for

dividing by two", and "dividing by two means".

In an alternate embodiment of the invention the patent examining process generates errors that have a correspondence to statutorily defined errors under § 112 and other sections. The output of the method and apparatus of checking a patent specification of the invention can produce computer generated office actions. Such office actions may be in the form generally used by the United States Patent and Trademark Office and thus serve as a quick and reliable method of spotting errors in a patent specification that are analogous, similar or identical to the ones described above. Such errors include reference character errors, missing parts of the specification, inconsistency within the specification with regard to elements and antecedent basis problems regarding the claims.

Three hypothetical examples of automated office actions are detailed immediately below. The first example is the case wherein the disclosure lacks reference characters. The second example covers the situation where part of the disclosure is missing. And the third example shows how a report can be generated using the method of the invention for vague and indefinite claims.

Example one shows the method of the invention verifying reference characters. A report is generated in the form of an official office action in accordance with regulations set down in the Manual of Patent Examining Procedure, latest revision November, 1989. If duplicate reference characters are used, or reference characters are missing, for different stated elements the generated office action may include the following paragraph in accordance with MPEP § 608.01 using form paragraph 7.29. In this hypothetical

example the reference character "10" was used in error for "110". An office action may be generated as follows.

The following is a quotation of the first paragraph of 35 U.S.C. 112: "The specification shall contain a written description of the invention and of the manner and process of making and using it, in such full clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

The disclosure is objected to because of the following informalities: In the specification on page 12, line 5 change reference character "10" to --110--. Appropriate correction is required.

Without the present invention, the examiner would manually find and note informalities. With the use of the present invention, the informalities are automatically reported and a notice is generated.

Example two shows the method of the checking for the existence of parts of the patent. In this hypothetical example, during automatic examination of the patent the Abstract section was discovered to be missing. Again a report is generated in the form of an official office action in accordance with regulations set down in the Manual of Patent Examining Procedure, latest revision November, 1989. These individual error reports may be automatically compiled to form an official office action. The following paragraph is taken according to form paragraph 6.12 as stated in the MPEP § 608.01(b).

This application does not contain an Abstract of the Disclosure as required by 37 CFR

1.72(b). An abstract on a separate sheet is required.

Example three shows the method of the invention for checking antecedents in the claims. The examiner would issue a rejection due to a vague or indefinite claim because of lack of antecedent basis for certain claim elements. This rejection may be automatically generated in accordance with MPEP § 706.03 form paragraph 7.34. After the form paragraph 7.34 the report generated by the method of the invention lists the claim elements found which have no antecedent basis as in the example shown below.

The following is a quotation of the first paragraph of 35 U.S.C. 112: "The specification shall contain a written description of the invention and of the manner and process of making and using it, in such full clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

Claim (1) rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which the applicant regards as the invention.

The following claim elements lack antecedent basis:

Claim 1, line 5, "the computer"; and

Claim 10, line 5, "the memory".

In such a way an examiner's office action covering inconsistent reference characters, missing parts of the disclosure and claims which are indefinite due to lack of antecedents may be generated

by the method of the invention.

Optionally, the invention, at the request of the user, can search for words that are considered non-standard usage in the patent law. These non-standard usage words, like the phrase "consisting of" within a claim generally detract from the quality of the patent. Other such words and their locations in the patent are listed below. It is recognized that some of the listed words may have proper applicability in limited, specific usages. In those special cases another list may be used, or the generated report may be modified by the user.

	Location	non-standard words
1.	Transition from claim preamble	having including wherein constituting consisting of consisting essentially
2.	Anywhere in the patent document	can would could will shall you
3.	In the Abstract of the Disclosure	comprises wherein means means for

The invention has been described herein in considerable detail in order to comply with the Patent Statutes and to provide those skilled in the art with the information needed to apply the novel principles and to construct and use such specialized components as are required. However, it is to be understood that the invention can be carried out by specifically different equipment and devices, and that various

modifications, both as to the equipment details and operating procedures, can be accomplished without departing from the scope of the invention itself.

What is claimed is:

CLAIMS

1. A patent disclosure checking apparatus (1000) for checking a patent disclosure in a digital form against a set of patent disclosure drafting rules (1700) comprising a processing means (10) for checking the patent disclosure against the set of patent disclosure drafting rules (1700).

2. The patent disclosure checking apparatus (1000) of claim 1 wherein the processing means (10) is selected from the group consisting of a digital computer, an analog computer, a microprocessor, boolean logic circuit, neural network circuit, an optical computer and a parallel processor.

3. The patent disclosure checking apparatus (1000) of claim 1 wherein the processing means (10) includes a means for reporting deviations which are a result of checking the patent disclosure against the set of patent disclosure drafting rules (1700).

4. The patent disclosure checking apparatus (1000) of claim 3 wherein the means for reporting deviations from the set of patent disclosure drafting rules (1700) reports deviations in a form comporting with an official office action.

5. An automated method for checking a document comprised of a serial group of characters, the automated method comprising the steps of:
   a. reading the serial group of characters (200-340);
   b. identifying a plurality of subgroups of serial characters which relate to each other

from within the serial group of characters (80); and

c. checking the plurality of subgroups for consistency of relation (440).

6. The automated method of claim 5 wherein the serial group of characters comprises a plurality of character strings, wherein the automated method further comprises the steps of:

a. reading the plurality of character strings wherein each character string comprises at least one character (250-300); and

b. identifying each one of the plurality of character strings as one of a plurality of string types wherein a reference character type is one of the plurality of string types describing a reference character (310-340).

7. The method of claim 6 further comprising the steps of:

a. allocating a unique location for each reference character (330);

b. selecting each reference character in turn and finding at least a first preceding word for each occurrence of each selected reference character (80);

c. identifying the at least first preceding word as an element of the selected reference character (380) and repeating steps b and c until all occurrences of all reference characters have been found;

d. selecting each reference character in turn and counting all identified elements (380) for each selected reference character; and

e. comparing all identified elements against

each other (440) for each selected reference character to determine usage inconsistencies.

8. The method of claim 7 further including the step of reporting identified elements which do not match the other identified elements for a selected reference character (100).

9. The method of claim 6 further including the step of reading (370) a keyword file (1100) wherein the keyword file (1100) includes a plurality of keywords wherein each one of the plurality of keywords includes a related pointing character which indicates that a number may precede or follow an adjacent keyword.

10. The method of claim 9 wherein the keywords comprise words selected from the group consisting of:

a n,
a b,
the n,
the b,
about n,
was n,
at n,
of n,
of b,
with b,
with n,
are n,
are b,
then n,
then b,

```
than n,
than b,
every n,
every b,
least n,
figure b,
figures b,
sigma b,
lb n,
lbs n,
mps n,
ms n,
meters n,
ft n,
volt n,
volts n,
duty n,
cycle n,
degree n,
degrees n,
bit n,
bits n,
hz n,
khz n,
vdc n,
vac n,
arcsec n,
arcseconds n,
to n,
and b,
from b,
. b,
claim b,
numeral b,
january b,
```

january n,
february b,
february n,
march b,
march n,
april b,
april n,
may n,
may b,
june n,
june b,
july n,
july b,
august n,
august b,
september n,
september b,
october n,
october b,
november n,
november b,
december n,
december b,
members b,
all n,
samples b,
table n; and

wherein a character "n" following a selected keyword represents "next" (520) and indicates that a number may follow the selected keyword and a character "b" following the selected keyword represents "back" (510) and indicates that a number may precede the selected keyword.

11. The method of claim 5 further including the step

of tokenizing subgroups of characters in the document after reading the serial group of characters.

12. The method of claim 5 further including the step of identifying patent subgroups of characters wherein the identified patent subgroups include,

i. a background section (700),

ii. a summary (710),

iii. a brief description of the drawings (720),

iv. a description of the preferred embodiment (730), and

v. a claims section (740).

13. The method of claim 12 wherein the step of identifying patent subgroups of characters further comprises the steps of matching patent subgroups of characters to predefined patent section headers (700-760).

14. The method of claim 13 wherein the predefined patent section headers include phrases selected from the group consisting of "BACKGROUND OF THE INVENTION", "SUMMARY OF THE INVENTION", "BRIEF DESCRIPTION OF THE DRAWINGS", "DESCRIPTION OF THE PREFERRED EMBODIMENT", and "I CLAIM".

15. The method of claim 12 further including the step of reserving a line count for each identified section so as to locate each section as needed for further processing.

16. The method of claim 5 wherein the step of reading the serial group of characters further includes the steps of:

a. filling an indexed string array with a currently selected character (230);
b. determining whether the currently selected character is punctuation (260);
c. if the currently selected character is determined to be punctuation then allocating the indexed string array to a state indicating punctuation (270);
d. terminating the indexed string array when a string delimiter is encountered (300);
e. determining whether the indexed string array is a reference character (320); and
f. if the indexed string array is determined to be a reference character then allocating the indexed string array to a state indicating a reference character (330), otherwise allocating the indexed string array to a state indicating a word (340).

17. The method of claim 16 wherein the step of allocating the indexed string array to a state indicating punctuation further includes the step of considering the special case where the currently selected character is a period and wherein a period is differentiated from a decimal point (350).

18. The method of claim 12 wherein the claims section includes at least one claim and wherein the method further comprises the steps of:
a. identifying each claim (800-834);
b. counting the at least one claim (120, 811-810);
c. further identifying each claim as a dependent or independent claim (924, 926);

and

d. for each identified dependent claim, further identifying a claim from which it depends (934, 938, 942, 946).

19. The method of claim 5 wherein the document is encoded in binary format.

20. The method of claim 5 wherein the step of checking the plurality of subgroups for consistency of relation further includes the step of reporting inconsistencies as errors in binary format.

21. The apparatus of claim 5 wherein the step of checking for consistency further includes the step of recognizing abbreviations and aliases.

22. The method of claim 9 wherein as an inconsistency is detected for a selected reference character a user of the method is requested to verify whether or not the serial group of characters causing the inconsistency is a new keyword (370) and if the serial group of characters causing the inconsistency is verified as the new keyword the keyword file (1100) is updated.

23. An automated method for checking a document comprising the steps of:

a. loading a keyword file (1100, 20);

b. reading the document for a set of predefined sections (40);

c. printing a portion of the document (50);

d. reading the document for reference characters and assigning a type designation

to each reference character (60);

e. selecting each reference character in turn and finding at least one preceding word for each occurrence of the selected reference character (370, 380);

f. identifying a first preceding word as an element of the selected reference character and repeating steps e and f until all occurrences of all reference characters have been found (80);

g. selecting each reference character in turn and counting all identified elements for each selected reference character (380); and

h. comparing all identified elements against each other for each selected reference character to determine a consistency of usage (90).

24. The automated method for checking a document of claim 23 further comprising the step of reporting at least one identified element which is not used consistently in relation to all identified elements for a selected reference character (100).

25. The automated method for checking a document of claim 23 further comprising the step of checking for words matching identified elements which are not proximate a reference character (110).

26. The method of claim 23 wherein as each error is detected for a selected reference character a user is requested to verify whether or not an indicated error is a keyword (370) and if the indicated error is verified as a keyword the

keyword file (1100) is updated.

27. An automated method of checking claims in a patent document which includes a claims section including at least one claim wherein the method comprises the steps of:
    a. identifying the at least one claim (800-834);
    b. counting the at least one claim (120, 810, 811);
    c. further identifying the at least one claim as a dependent or independent claim (924, 926); and
    d. for each identified dependent claim, further identifying a claim from which it depends (934, 938, 942, 946).

28. An automated method for checking a consistency of usage of figure references in a patent specification including a brief description of the drawings section wherein at least one drawing is referred to by at least one figure reference, the method comprising the steps of:
    a. locating the brief description of the drawing section (1104);
    b. reading in the at least one figure reference wherein each figure reference includes at least one figure number and at least one figure description in the brief description of the drawings section (1106, 1108);
    c. reading the patent specification exclusive of the brief description of the drawings section for specification figure numbers and specification figure descriptions (1112-1116); and

d. checking a consistency of usage of the specification figure numbers and figure descriptions against the brief description of the drawings section's at least one figure number and at least one figure description (1118).

29. The automated method of claim 28 further comprising the step of reporting inconsistencies of usage (1120) in a form of a patent and trademark office official communication.

30. An automated method of finding non-standard usage of words within a patent specification, the patent specification being organized into a plurality of parts, each part comprising a plurality of words, the method comprising the steps of:

a. reading in a list of non-standard words and their associated patent part;

b. reading each word in the patent specification and noting a part of the patent specification within which each word occurs;

c. for each word read determining whether it is on the list of non-standard words for the patent specification part being read; and

d. reporting an error if a read word is determined to be on the list of non-standard words.

31. A method of automatically correcting a patent specification encoded in a computer, the patent specification encoded in a computer having at least one drafting error, comprising the steps

of:

a. locating at least one drafting error of the patent specification encoded; and

b. automatically correcting the at least one drafting error directly into the patent specification encoded in a computer.

32. An automated method of checking usage consistency of drawing elements in a patent specification having at least one drawing with at least one reference character in the patent specification comprising the steps of:

a. reading in a list of drawing elements comprising at least one drawing element and an associated reference character for the at least one drawing element (1104-1108);

b. checking the list of drawing elements against each use of the at least one reference character in the patent specification for consistency of usage (1112-1118); and

c. reporting any usage inconsistencies including omissions of reference characters (1120).

33. An automated method of counting at least one claim, wherein the at least one claim comprises a plurality of words, the automated method of counting claims comprising the steps of:

a. locating a lead claim with a lead claim number (904);

b. setting a claim count to one if the lead claim is found (906);

c. checking whether the words after the lead claim number are consistent with claim

drafting form (908, 910);

d. checking for a next claim with a next claim number (912);

e. incrementing a claim count if the next claim is found (912);

f. checking whether the words after the next claim number are consistent with claim drafting form (908, 910); and

g. repeating steps d through f until all claims have been counted.

34. The automated method of counting claims according to claim 33 further comprising the step of reporting claim numbers that do not follow a predetermined sequence (914).

35. The method of claim 33 wherein the step of locating a lead claim further comprises the step of obtaining the location of the lead claim by searching for claim language (804).

36. An automated method of checking for antecedent basis in at least one claim, the automated method of checking comprising the steps of:

a. searching for a claim element within the at least one claim (810);

b. determining whether the claim element is a claim antecedent and if the claim element is a claim antecedent adding the claim antecedent to a list of antecedents (820);

c. determining whether the claim element needs a claim antecedent (820), and if the claim element needs a claim antecedent, then checking the list of antecedents for the claim element (822-834); and

d. repeating steps b and c until all claim elements have been checked.

37. The automated method of checking for antecedent basis of claim 36 further comprising the step of checking whether a claim element to be added to the list of antecedents is already in the list of antecedents (822) and if the claim element is already in the list of antecedents, reporting an error (830).

38. The automated method of checking for antecedent basis of claim 36, wherein the at least one claim is a dependent claim and the dependent claim has a list of antecedents for claim elements of the dependent claim wherein the automated method further comprises the step of checking the list of antecedents for claim elements of the dependent claim (824-832).

39. The automated method of checking for antecedent basis of claim 36 further comprising the step of reporting an error if an antecedent is not found in the antecedent list (834).

40. The automated method of checking for antecedent basis of claim 36 wherein the list of antecedents comprises a linked list.

41. The method of checking claims for antecedents according to claim 36 further including the step of reporting claim element antecedents not referenced by other claim elements.

42. The method of claim 36 further comprising the

step of reporting any claim elements lacking antecedent basis (830, 834) in a report generated according to a set of national patent office guidelines.

43. The method of claim 42 where the set of national patent office guidelines comprise United States Patent and Trademark Office guidelines.

44. An automated method for finding claims in a patent document having a claims section including a beginning of the claims section, the method comprising the steps of:

a. loading a current claim number from the beginning of the claims section (904);

b. scanning from the beginning of the claims section (906);

c. loading a next word from the claims section (908);

d. if the next word is null, then noting end of claims and stopping (916);

e. checking the next word against the current claim number (910);

f. if the next word is not similar to the current claim number, then logically repeating step c and following steps;

g. if the next word is similar to the current claim number, then change the current claim number (912); and

h. associating the current claim number with a current location (914).

45. The method of claim 44 further including the step of noting a highest claim number before stopping (916).

46. An automated method for finding dependent claims in a patent document having a claims section wherein the claims section begins at a predetermined location, the method comprising the steps of:

a. loading a first claim as a current claim (922);

b. reading the current claim to find references to claims (924);

c. if there are no references to claims in the current claim, then indicating the current claim as an independent claim (926);

d. if there are references to claims in the current claim, then determining whether the current claim refers to itself (934);

e. if there are references to claims in the current claim, then determining whether the current claim refers to a nonexistent claim (938);

f. if there are references to claims in the current claim, then determining whether the current claim refers to a higher numbered claim (942); and

g. loading a next claim as the current claim (928-930) and logically repeating steps b through g.

47. The method of claim 46 further including the step of reporting an error if the current claim refers to itself (936).

48. The method of claim 46 further including the step of reporting an error if the current claim refers to a nonexistent claim (940).

49. The method of claim 46 further including the step of reporting an error if the current claim refers to a higher numbered claim (944).

50. The method of claim 46 further including the step of reporting a dependency structure of all claims (932).

51. An automated method for finding claim elements (810A) in a patent document having a claims section wherein the claims section begins at a first word of the claims section and the claims section includes at least one predetermined trigger word, the method comprising the steps of:

a. loading the first word of the claims section as a current word (972);

b. checking whether the current word is a trigger word (974);

c. if the current word is not a trigger word, then loading a next word in the claims section as the current word (986) and repeating step b;

d. reading forward from the trigger word to a predetermined claim word or to an end of the claims section whichever occurs first to determine a claim element (976-978);

e. if no claim word or end of the claims section is found, then reporting an error (980);

f. loading the next word in the claims section as the current word (986); and

g. repeating steps b through g if the current word is not null (988).

52. The method of claim 51 wherein the at least one

predetermined trigger word is selected from the group consisting of:

a,
an,
means for,
at least one,
the, and
said.

53. The method of claim 52 wherein the step of identifying all claim elements for each claim (976-978) further includes the step of looking backward at a word string preceding the "means for" trigger word for that claim and if a word preceding the word string is null or a number or a claim word then allocating the word string as a claim antecedent (982).

54. The method of claim 51 wherein a plurality of claim antecedents are generated for each claim element, and wherein the plurality of claim antecedents are synonymous with the claim element (984).

55. The method of claim 51 wherein the predetermined claim word is selected from the group consisting of:

"comprising",
"containing",
"consisting",
",",
".",
";",
":",
"(",

"comprises",
"includes",
"having",
"has",
"for",
"is",
"wherein",
"further",
"are",
"adapted",
"which",
"and",
"such",
"in",
"along",
"around",
"below",
"above",
"over",
"under",
"beneath",
"after",
"before",
"means"; and
"consists".

56. The method of claim 51 wherein the step of identifying all claim elements for each claim (976-978) further includes the steps of:

a. in a case where no antecedent delimiter is found and no "means for" phrase is found, looking for a phrase "at least one" by checking the current word string against "at" and looking for "least" and "one" following "at";

b. reading forward from "at least one" or "at

least";

c. finding a word "means" then looking for a word "for" to be a word string immediately following "means";

d. if "means for" is found then reading "means for" until a current word looking forward is null or a number or a claim word; and

e. allocating the forward word string as an antecedent (982).

57. The method of claim 36 wherein the step of checking the list of antecedents (822-834) for the claim element further comprises the steps of:

a. checking the claim element on a word by word basis against each antecedent in the list of antecedents (822, 824, 832);

b. determining if any antecedent in the list of antecedents does not match the claim element (822, 824, 832).

58. A method of checking claim antecedents in at least one claim comprising the steps of:

a. for the at least one claim identifying at least one antecedent claim element (820);

b. storing the at least one antecedent claim element in an antecedent data structure which points to all antecedents arranged by claim number (826);

c. locating at least one needy claim element (820) which requires an antecedent claim element;

d. testing the at least one needy claim element against the antecedent data structure (824); and

e. determining whether or not the needy claim

element relates consistently to the at least one antecedent claim element (832).

59. The method of claim 36 wherein the at least one claim comprises a plurality of words starting with a first word, and wherein the step of finding a claim element further comprises the steps of:

a. loading the first word of the at least one claim as a current word (972);

b. checking whether the current word is a trigger word (974);

c. if the current word is not a trigger word, then loading a next word in the at least one claim as the current word and repeating step b (986);

d. reading forward from the trigger word to a predetermined claim word or to the end of the at least one claim whichever occurs first to determine a claim element (978);

e. if no claim word or end of the at least one claim is found, then reporting an error (980);

f. loading the next word in the at least one claim as the current word (986); and

g. repeating steps b through g if the current word is not null (988).

60. An automated method of generating a national patent office action following a set of national patent office action guidelines.

61. The method of claim 57 wherein the set of national patent office action guidelines are United States Patent and Trademark Office

guidelines.

62. The apparatus of claim 1 wherein the patent disclosure is comprised of at least one document element, and wherein the digital form includes a document element structure (500) comprising:

a. means for pointing to a preceding document element (510) associated with the at least one document element (500);

b. means for pointing to a next element (520) associated with the at least one document element (500);

c. means for pointing to a string storage location (530) for the document element structure associated with the at least one document element (500); and

d. means for pointing to a document element type (540) associated with the at least one document element (500).

63. The apparatus of claim 59 wherein the digital form includes a reference character element (600) structure comprising:

a. means for pointing to a next reference element (610);

b. means for pointing to a document element (630) for the document element associated with the reference character element (600); and

c. means for pointing to a numerical value (620).

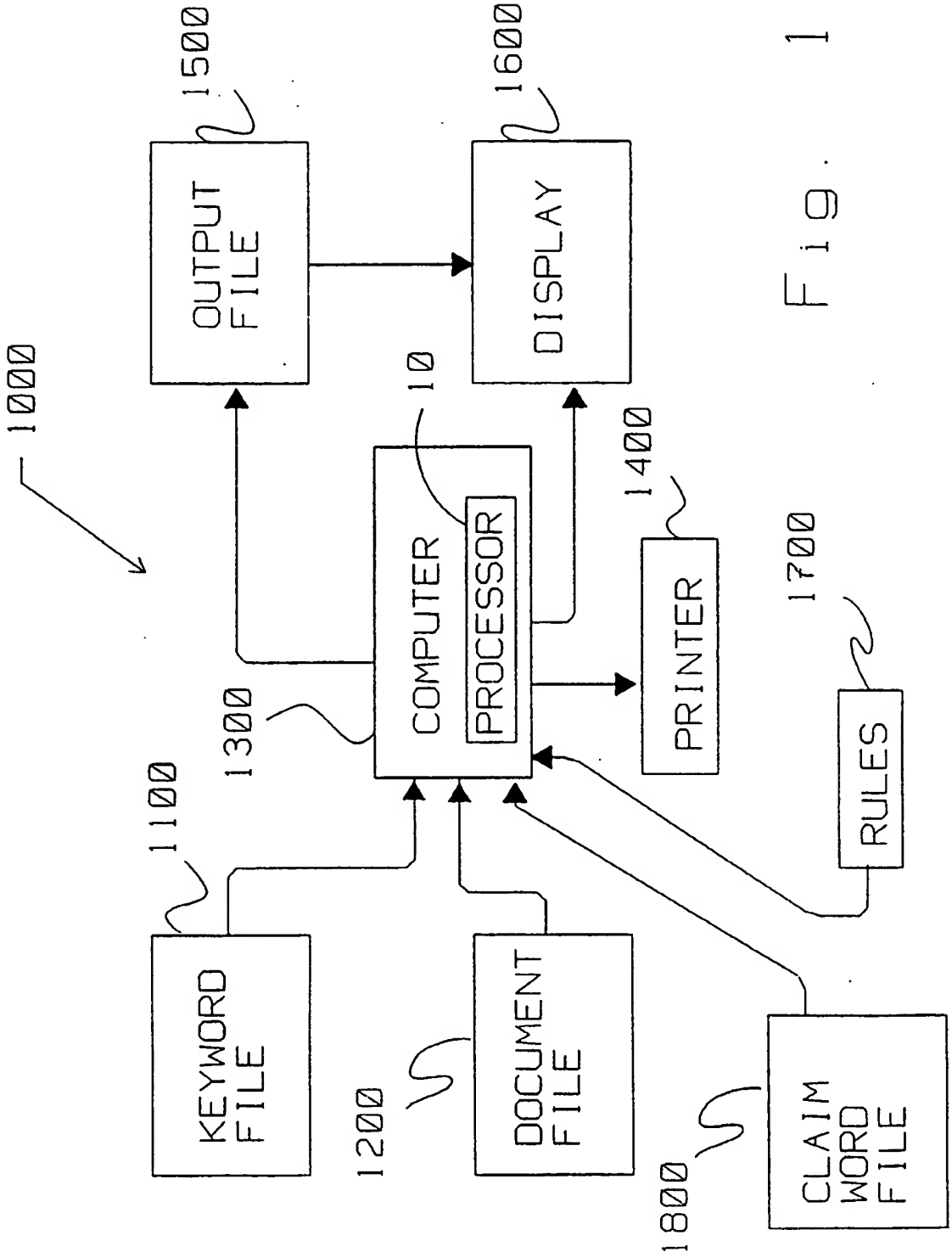

1000
OUTPUT FILE
1500
DISPLAY
1600
10
COMPUTER
PROCESSOR
1300
1400
PRINTER
1700
RULES
1100
KEYWORD FILE
1200
DOCUMENT FILE
1800
CLAIM WORD FILE

Fig. 1

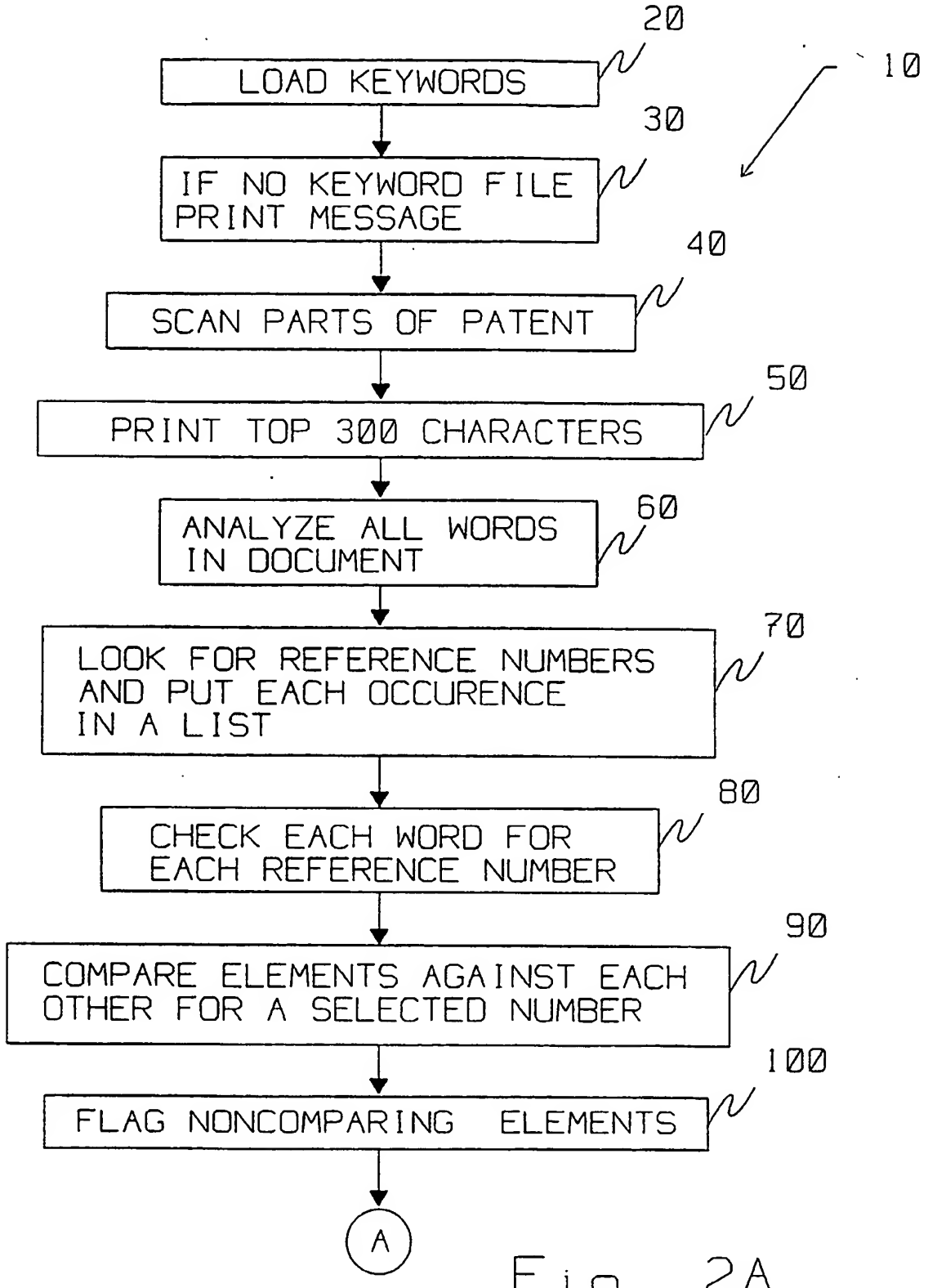

10
20
LOAD KEYWORDS
30
IF NO KEYWORD FILE PRINT MESSAGE
40
SCAN PARTS OF PATENT
50
PRINT TOP 300 CHARACTERS
60
ANALYZE ALL WORDS IN DOCUMENT
70
LOOK FOR REFERENCE NUMBERS AND PUT EACH OCCURENCE IN A LIST
80
CHECK EACH WORD FOR EACH REFERENCE NUMBER
90
COMPARE ELEMENTS AGAINST EACH OTHER FOR A SELECTED NUMBER
100
FLAG NONCOMPARING ELEMENTS
A

Fig. 2A

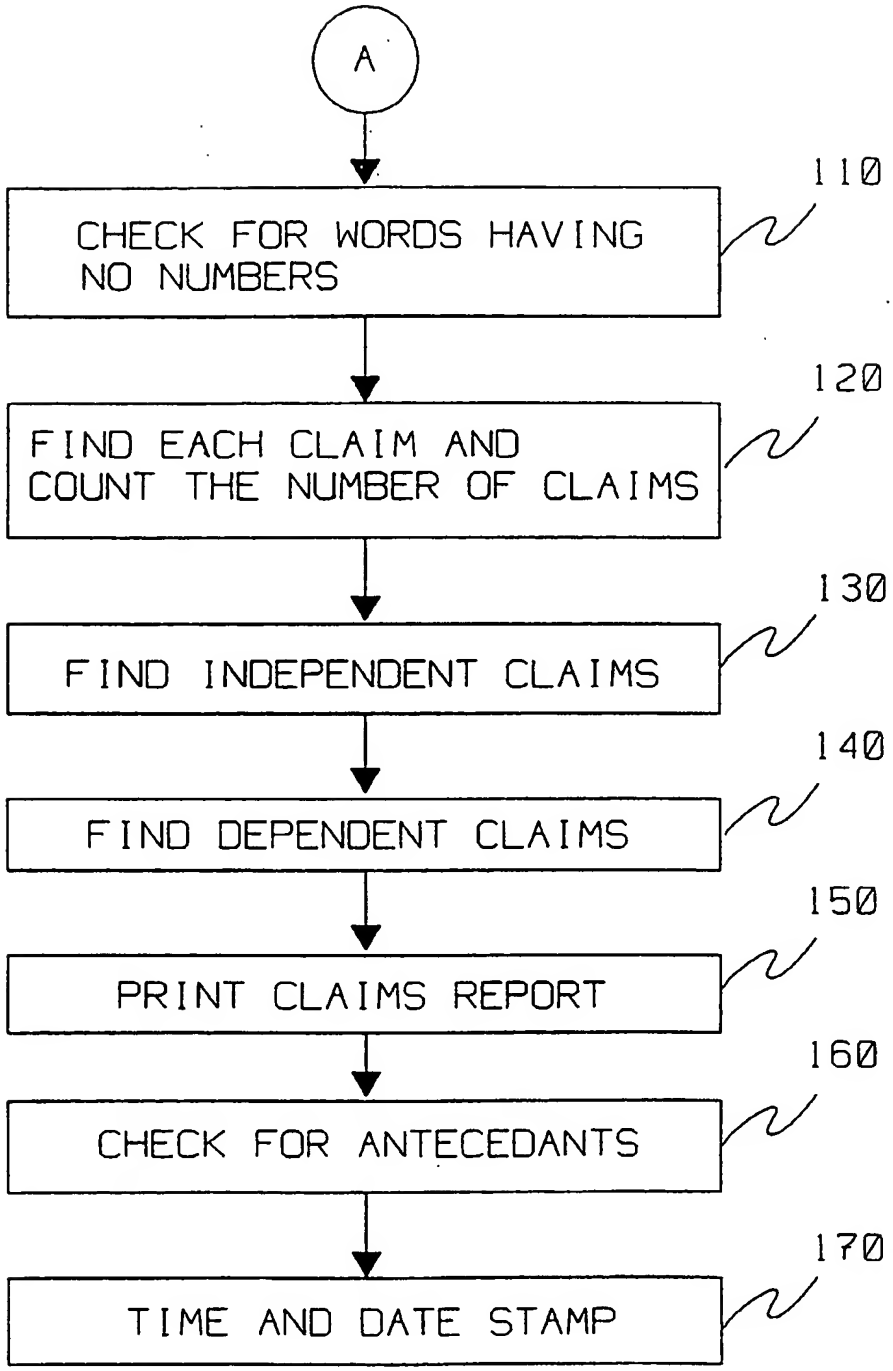

A
CHECK FOR WORDS HAVING NO NUMBERS
110
FIND EACH CLAIM AND COUNT THE NUMBER OF CLAIMS
120
FIND INDEPENDENT CLAIMS
130
FIND DEPENDENT CLAIMS
140
PRINT CLAIMS REPORT
150
CHECK FOR ANTECEDANTS
160
TIME AND DATE STAMP
170

Fig. 2B

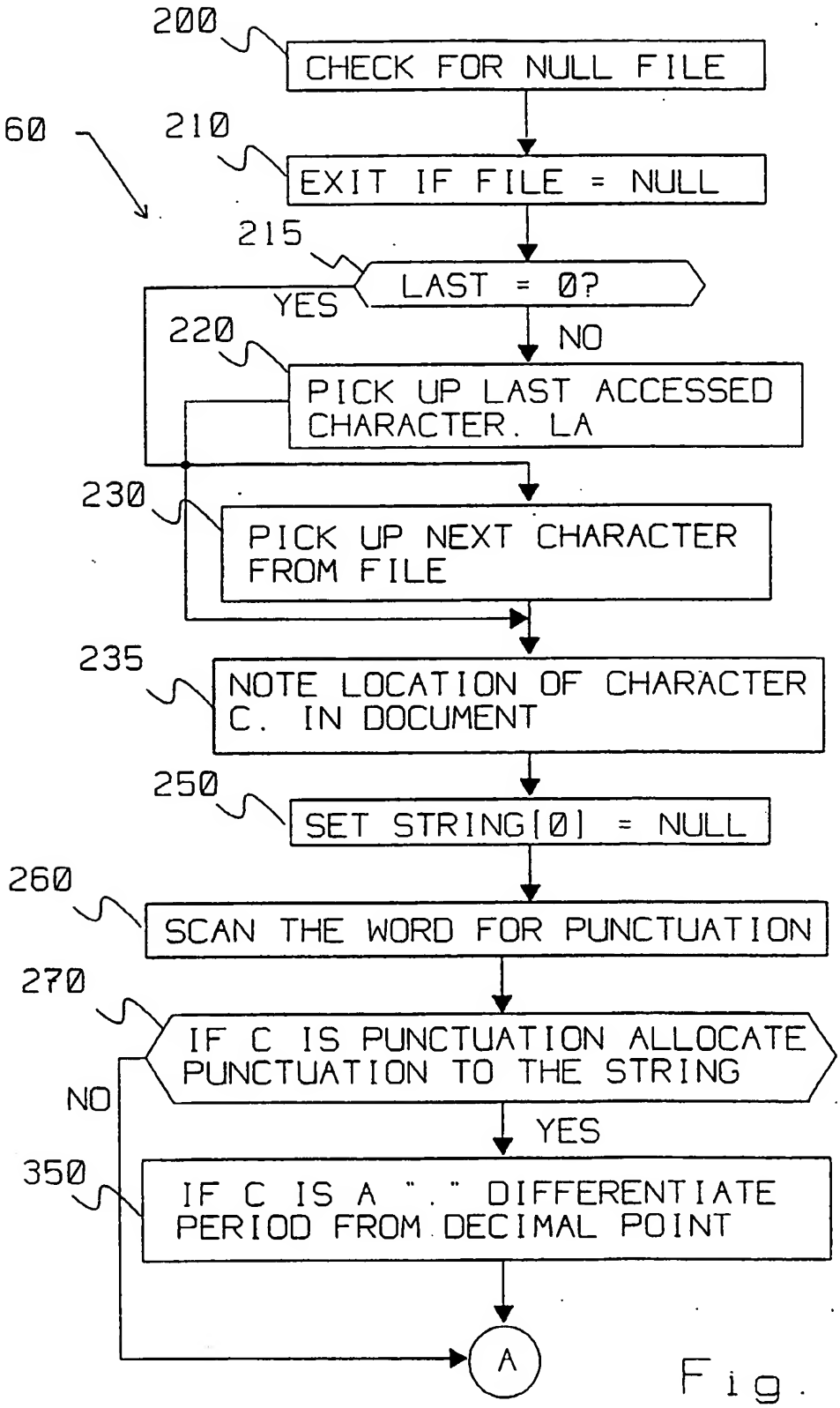

200
CHECK FOR NULL FILE
60
210
EXIT IF FILE = NULL
215
LAST = 0?
YES
NO
220
PICK UP LAST ACCESSED CHARACTER. LA
230
PICK UP NEXT CHARACTER FROM FILE
235
NOTE LOCATION OF CHARACTER C. IN DOCUMENT
250
SET STRING[0] = NULL
260
SCAN THE WORD FOR PUNCTUATION
270
IF C IS PUNCTUATION ALLOCATE PUNCTUATION TO THE STRING
NO
YES
350
IF C IS A "." DIFFERENTIATE PERIOD FROM DECIMAL POINT
A

Fig. 3A

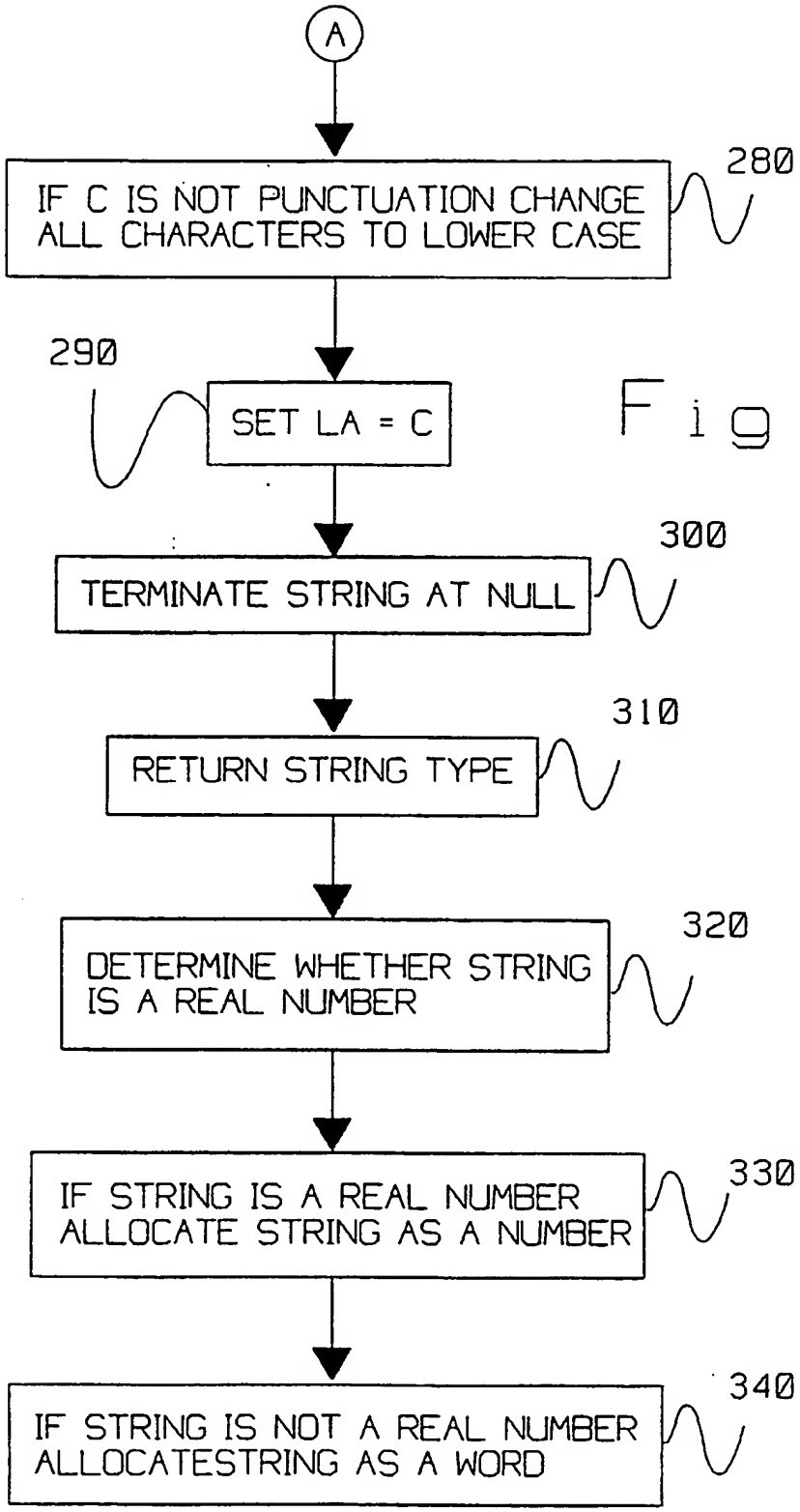
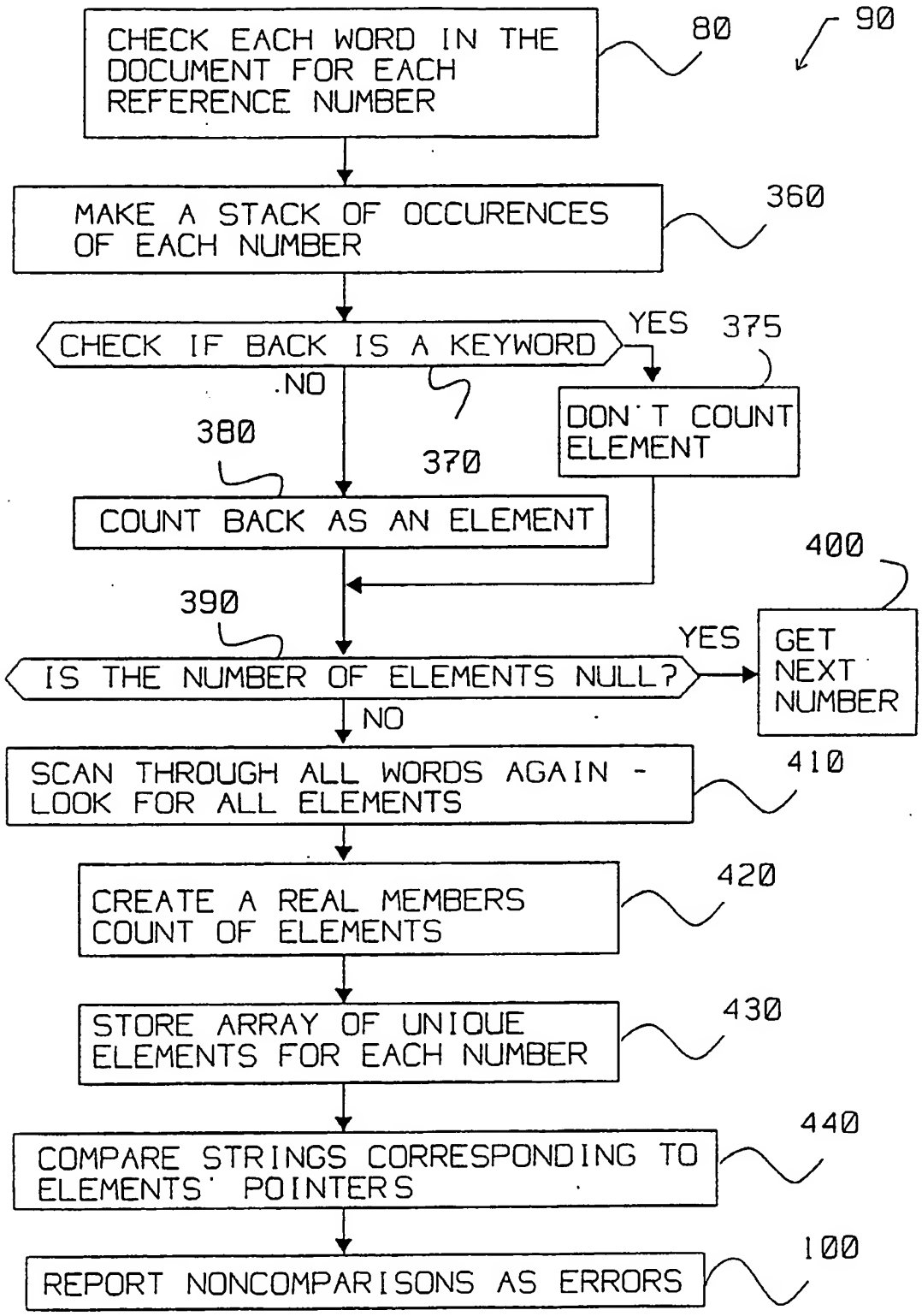

A
IF C IS NOT PUNCTUATION CHANGE ALL CHARACTERS TO LOWER CASE
280
290
SET LA = C
TERMINATE STRING AT NULL
300
RETURN STRING TYPE
310
DETERMINE WHETHER STRING IS A REAL NUMBER
320
IF STRING IS A REAL NUMBER ALLOCATE STRING AS A NUMBER
330
IF STRING IS NOT A REAL NUMBER ALLOCATESTRING AS A WORD
340

Fig. 3B

Fig. 4


CHECK EACH WORD IN THE DOCUMENT FOR EACH REFERENCE NUMBER
80
90
MAKE A STACK OF OCCURENCES OF EACH NUMBER
360
CHECK IF BACK IS A KEYWORD
YES
375
NO
380
370
DON'T COUNT ELEMENT
COUNT BACK AS AN ELEMENT
400
390
YES
GET NEXT NUMBER
IS THE NUMBER OF ELEMENTS NULL?
NO
SCAN THROUGH ALL WORDS AGAIN - LOOK FOR ALL ELEMENTS
410
CREATE A REAL MEMBERS COUNT OF ELEMENTS
420
STORE ARRAY OF UNIQUE ELEMENTS FOR EACH NUMBER
430
COMPARE STRINGS CORRESPONDING TO ELEMENTS' POINTERS
440
REPORT NONCOMPARISONS AS ERRORS
100

WORD STRUCTURE

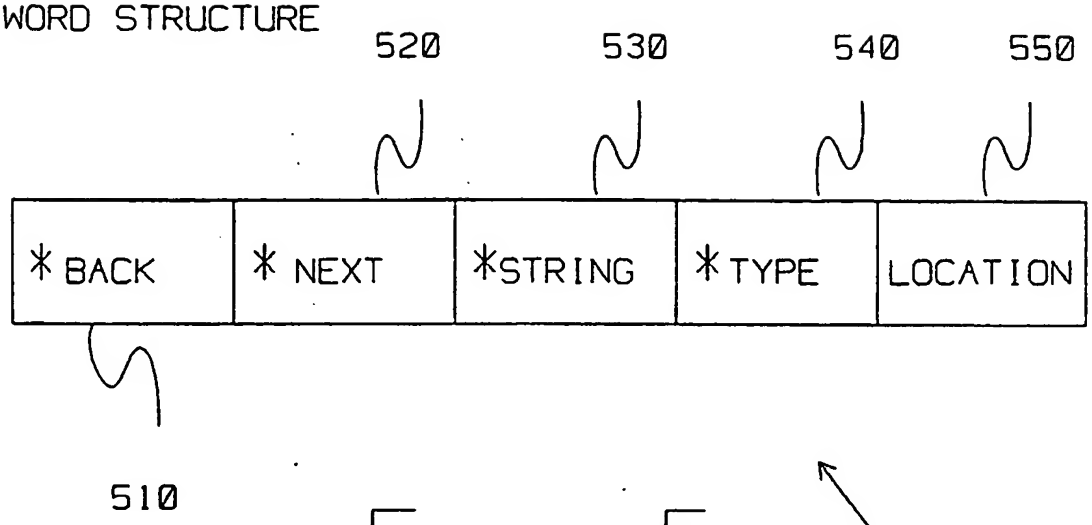


Fig. 5

NUMBER STRUCTURE

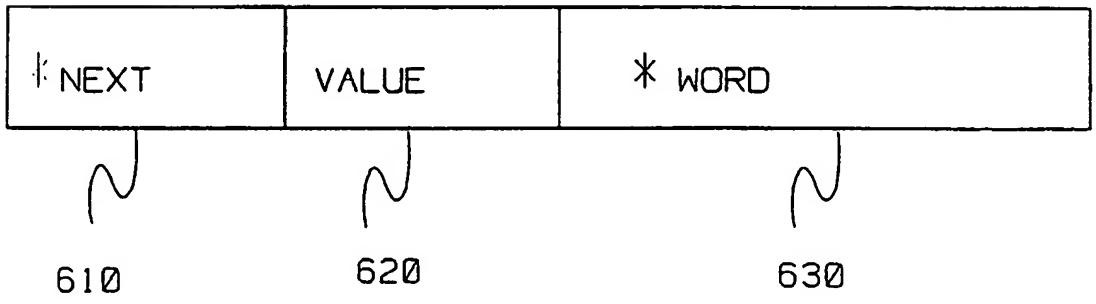


Fig. 6

Fig. 7

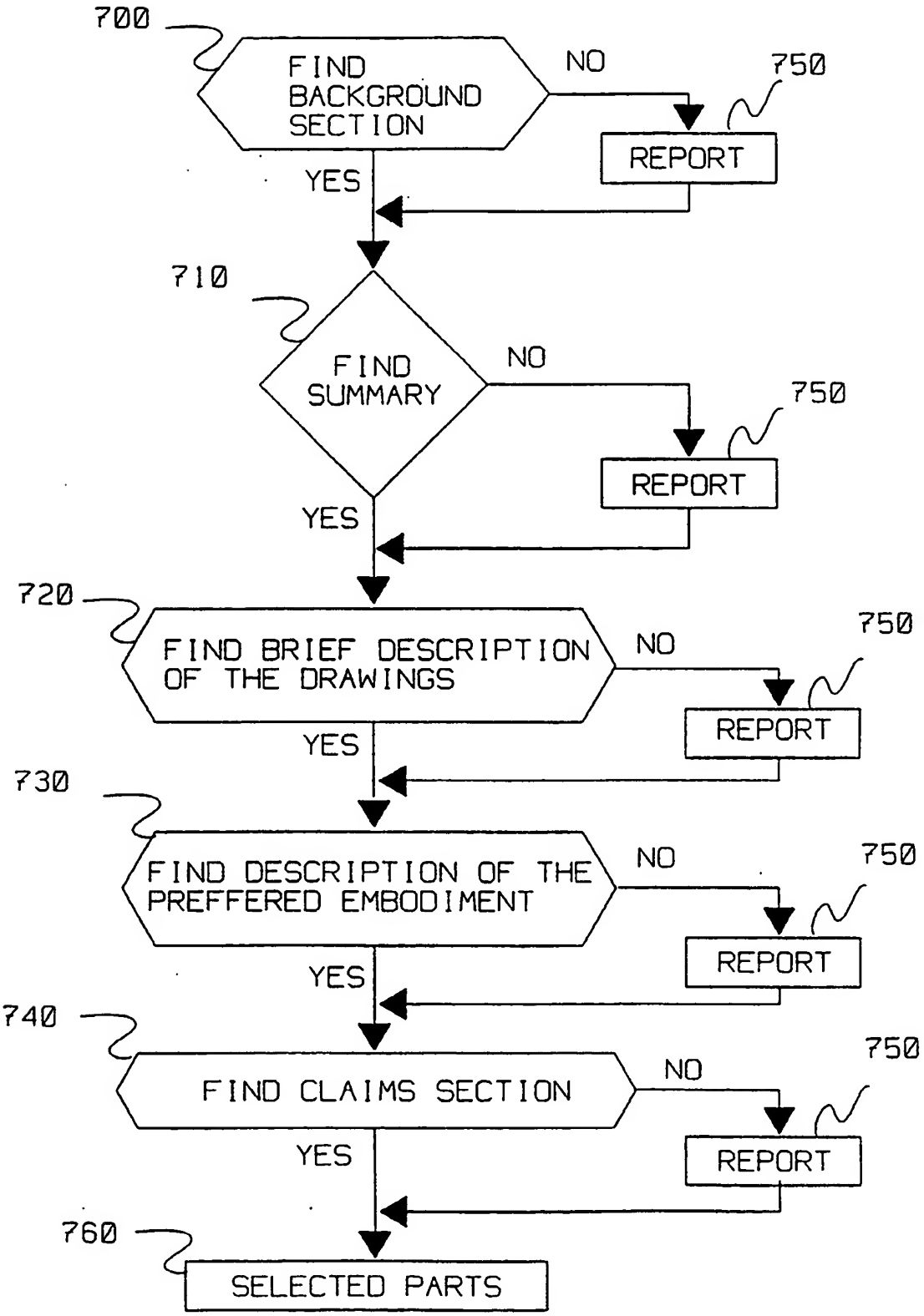

700
FIND BACKGROUND SECTION
NO
750
REPORT
YES
710
FIND SUMMARY
NO
750
REPORT
YES
720
FIND BRIEF DESCRIPTION OF THE DRAWINGS
NO
750
REPORT
YES
730
FIND DESCRIPTION OF THE PREFFERED EMBODIMENT
NO
750
REPORT
YES
740
FIND CLAIMS SECTION
NO
750
REPORT
YES
760
SELECTED PARTS

Fig. 8

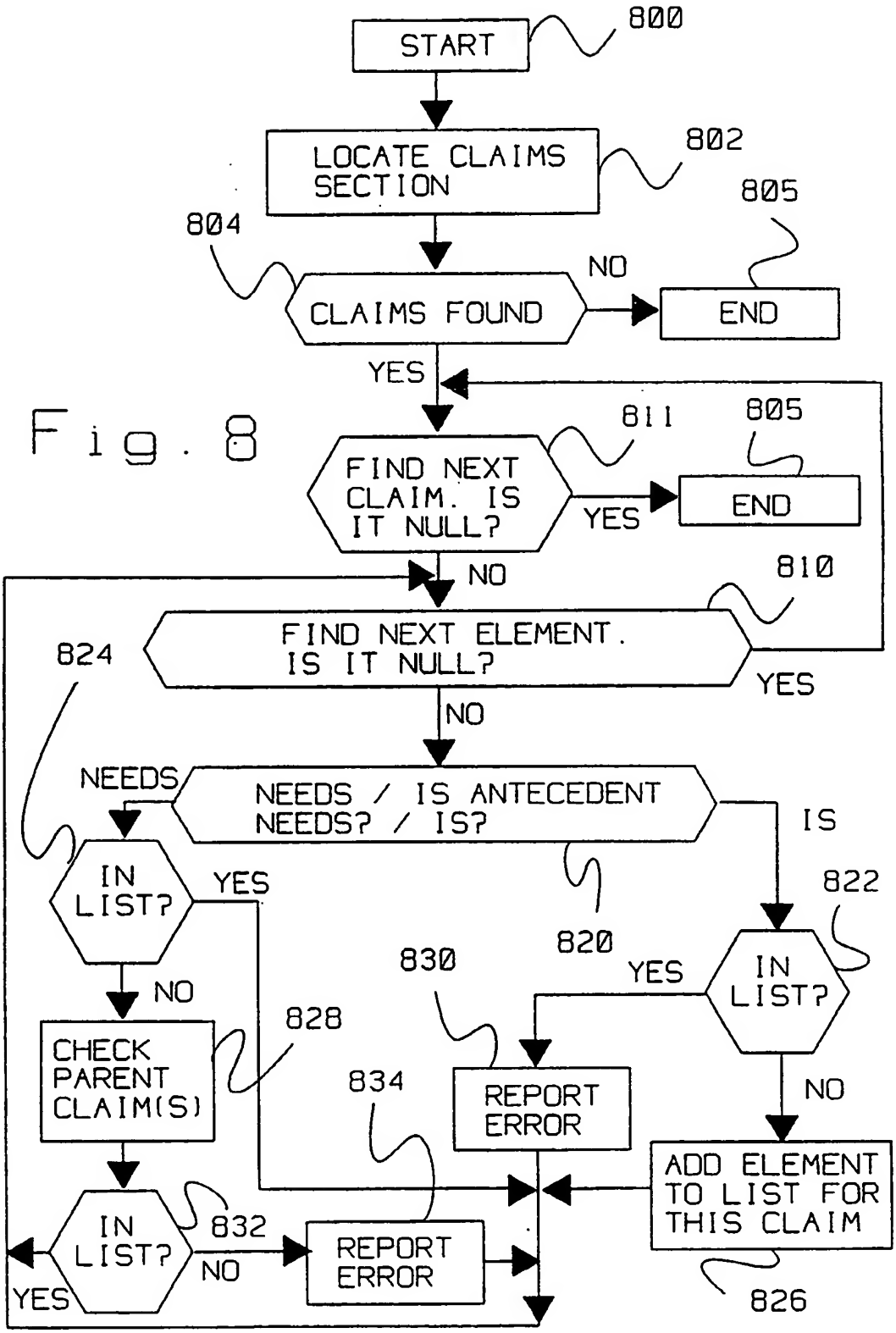

START
800
LOCATE CLAIMS SECTION
802
804
CLAIMS FOUND
NO
805
END
YES
FIND NEXT CLAIM. IS IT NULL?
811
YES
805
END
NO
FIND NEXT ELEMENT. IS IT NULL?
810
824
YES
NO
NEEDS
NEEDS / IS ANTECEDENT NEEDS? / IS?
IS
IN LIST?
YES
820
822
830
YES
IN LIST?
NO
CHECK PARENT CLAIM(S)
828
834
REPORT ERROR
NO
ADD ELEMENT TO LIST FOR THIS CLAIM
IN LIST?
832
NO
REPORT ERROR
YES
826

Fig. 9

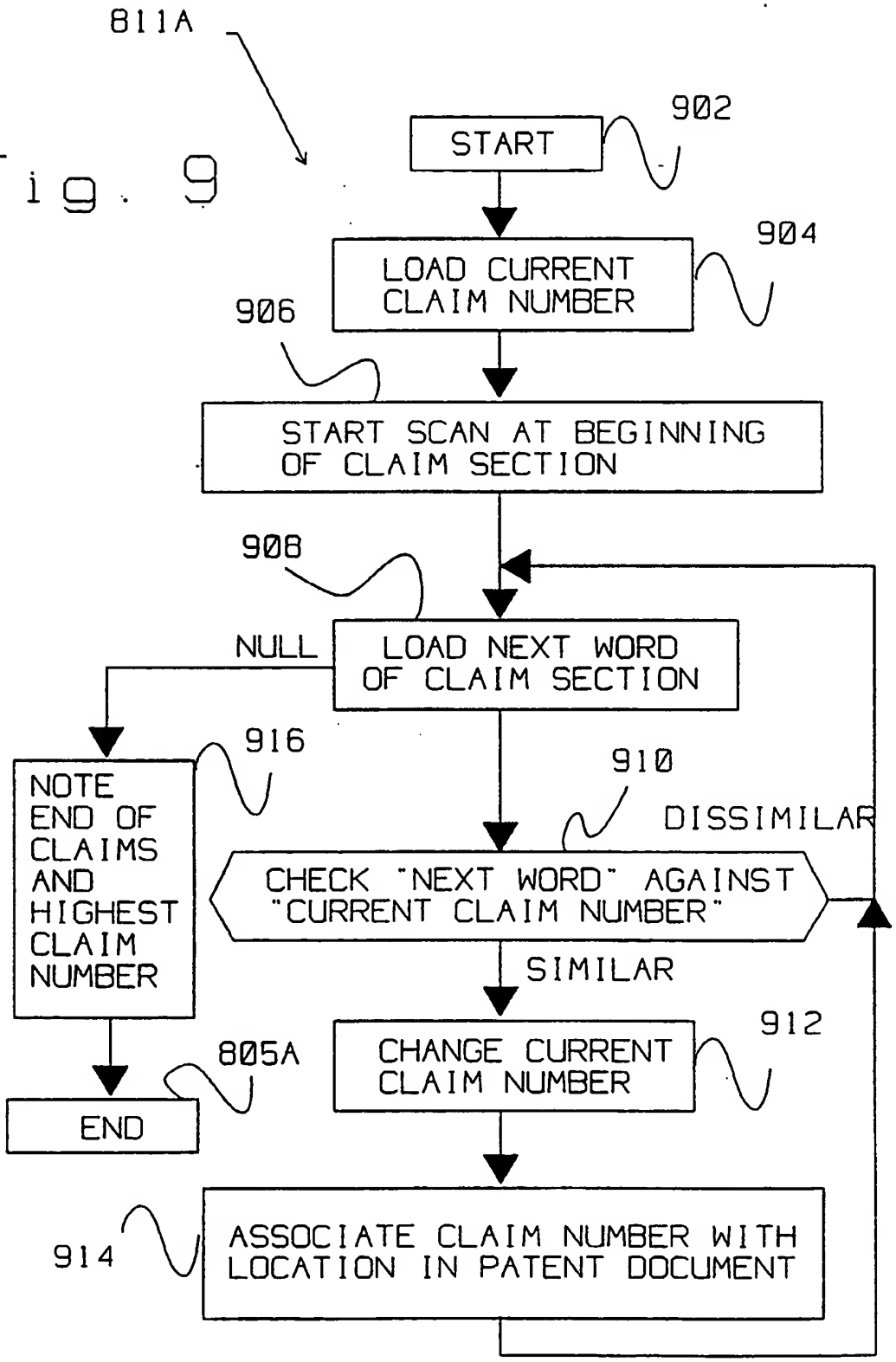

811A
902
START
904
LOAD CURRENT CLAIM NUMBER
906
START SCAN AT BEGINNING OF CLAIM SECTION
908
NULL
LOAD NEXT WORD OF CLAIM SECTION
916
NOTE END OF CLAIMS AND HIGHEST CLAIM NUMBER
910
DISSIMILAR
CHECK "NEXT WORD" AGAINST "CURRENT CLAIM NUMBER"
SIMILAR
912
CHANGE CURRENT CLAIM NUMBER
805A
END
914
ASSOCIATE CLAIM NUMBER WITH LOCATION IN PATENT DOCUMENT

Fig. 10

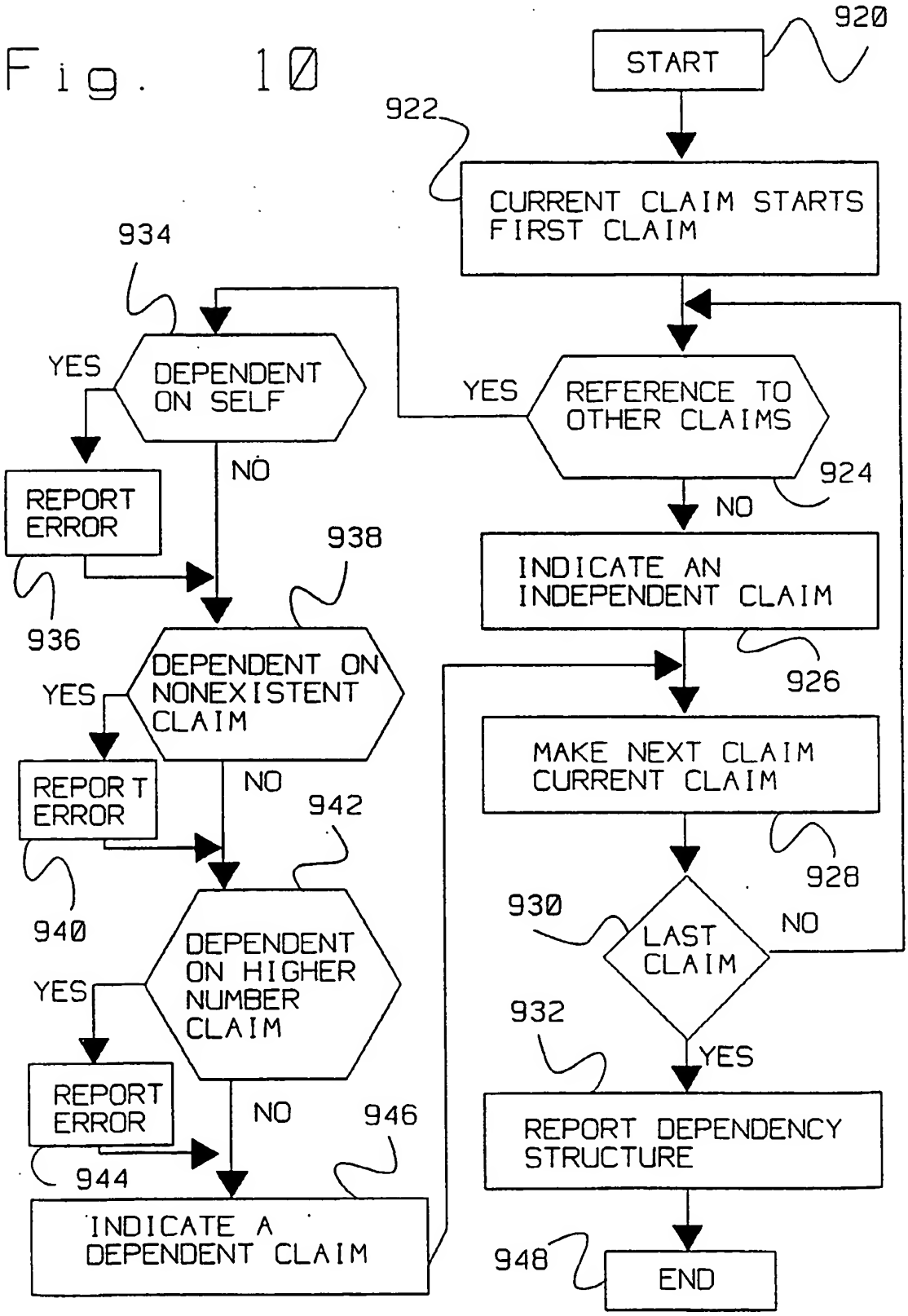
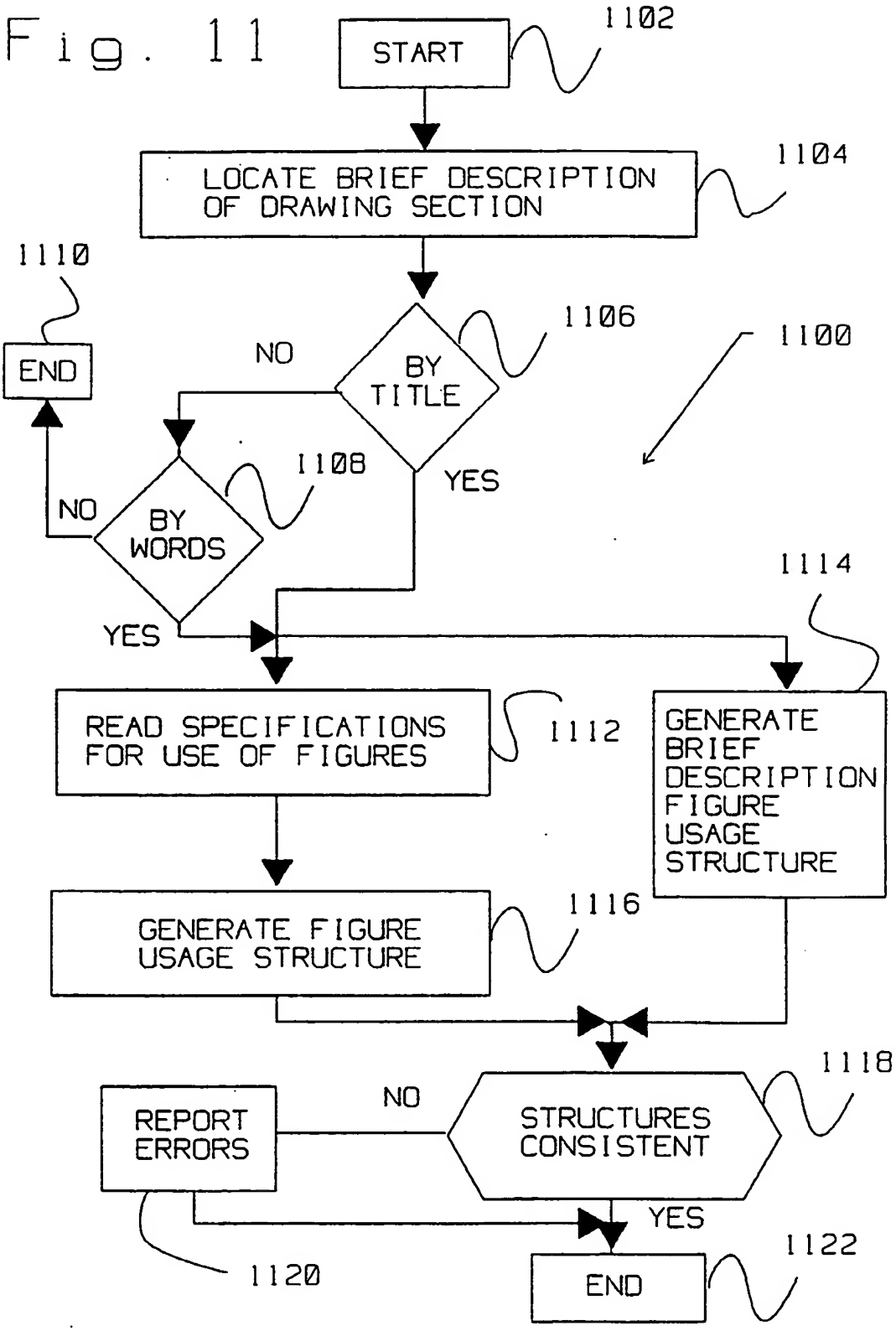

920
START
922
CURRENT CLAIM STARTS FIRST CLAIM
934
YES
DEPENDENT ON SELF
NO
REPORT ERROR
936
938
DEPENDENT ON NONEXISTENT CLAIM
YES
NO
REPORT ERROR
940
942
DEPENDENT ON HIGHER NUMBER CLAIM
YES
NO
REPORT ERROR
944
946
INDICATE A DEPENDENT CLAIM
YES
REFERENCE TO OTHER CLAIMS
924
NO
INDICATE AN INDEPENDENT CLAIM
926
MAKE NEXT CLAIM CURRENT CLAIM
928
930
LAST CLAIM
NO
932
YES
REPORT DEPENDENCY STRUCTURE
948
END

Fig. 11


START
1102
LOCATE BRIEF DESCRIPTION OF DRAWING SECTION
1104
1100
BY TITLE
1106
NO
YES
END
1110
BY WORDS
1108
NO
YES
1114
READ SPECIFICATIONS FOR USE OF FIGURES
1112
GENERATE BRIEF DESCRIPTION FIGURE USAGE STRUCTURE
GENERATE FIGURE USAGE STRUCTURE
1116
STRUCTURES CONSISTENT
1118
NO
REPORT ERRORS
YES
1120
END
1122

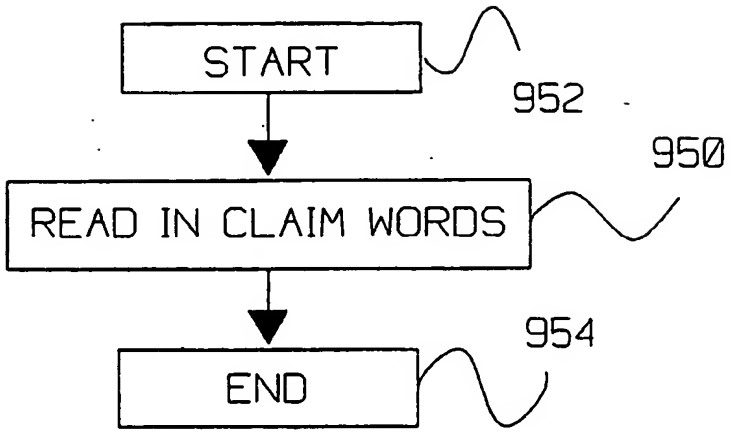

START
952
950
READ IN CLAIM WORDS
954
END

Fig. 12

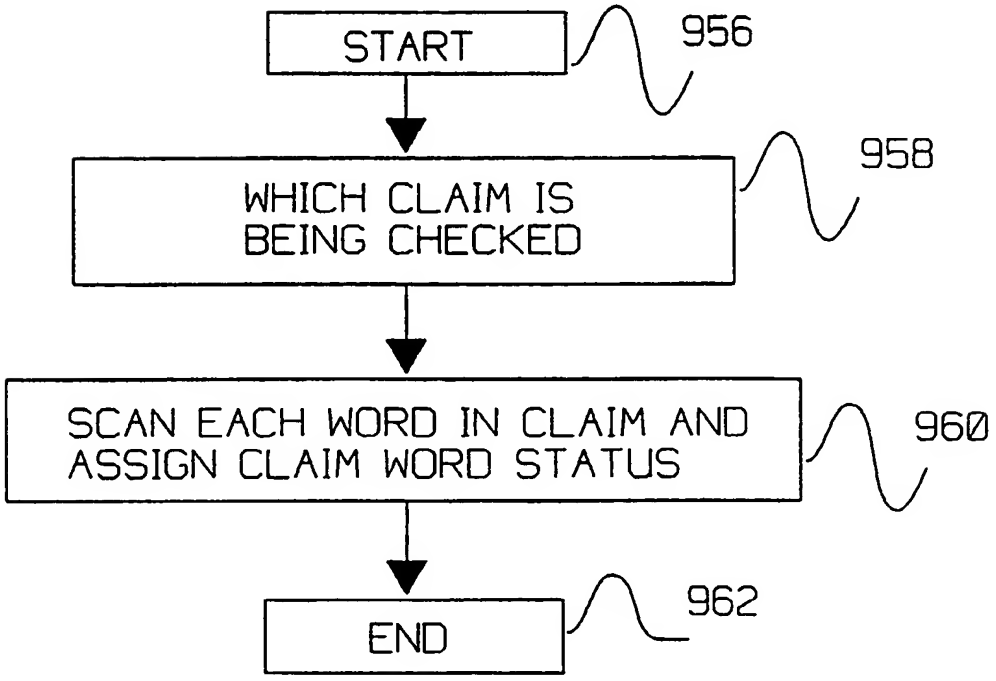

START
956
958
WHICH CLAIM IS
BEING CHECKED
SCAN EACH WORD IN CLAIM AND
ASSIGN CLAIM WORD STATUS
960
END
962

Fig. 13

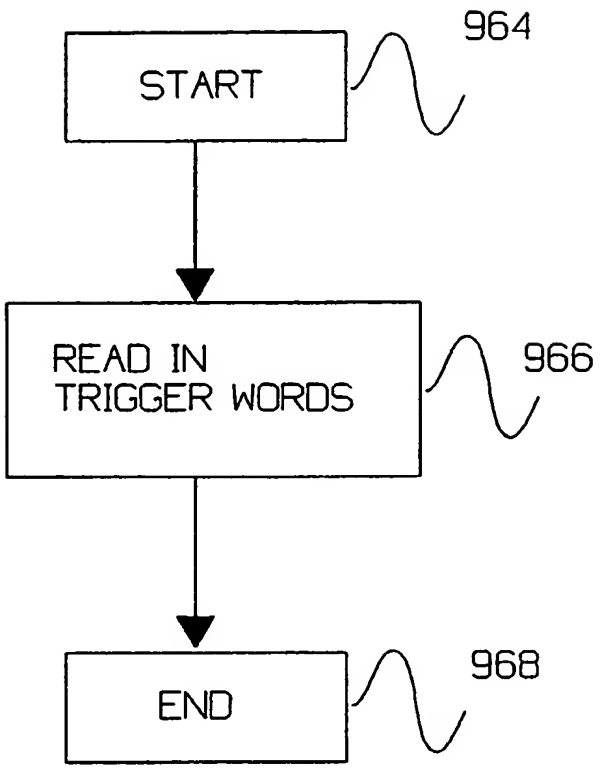


Fig. 14

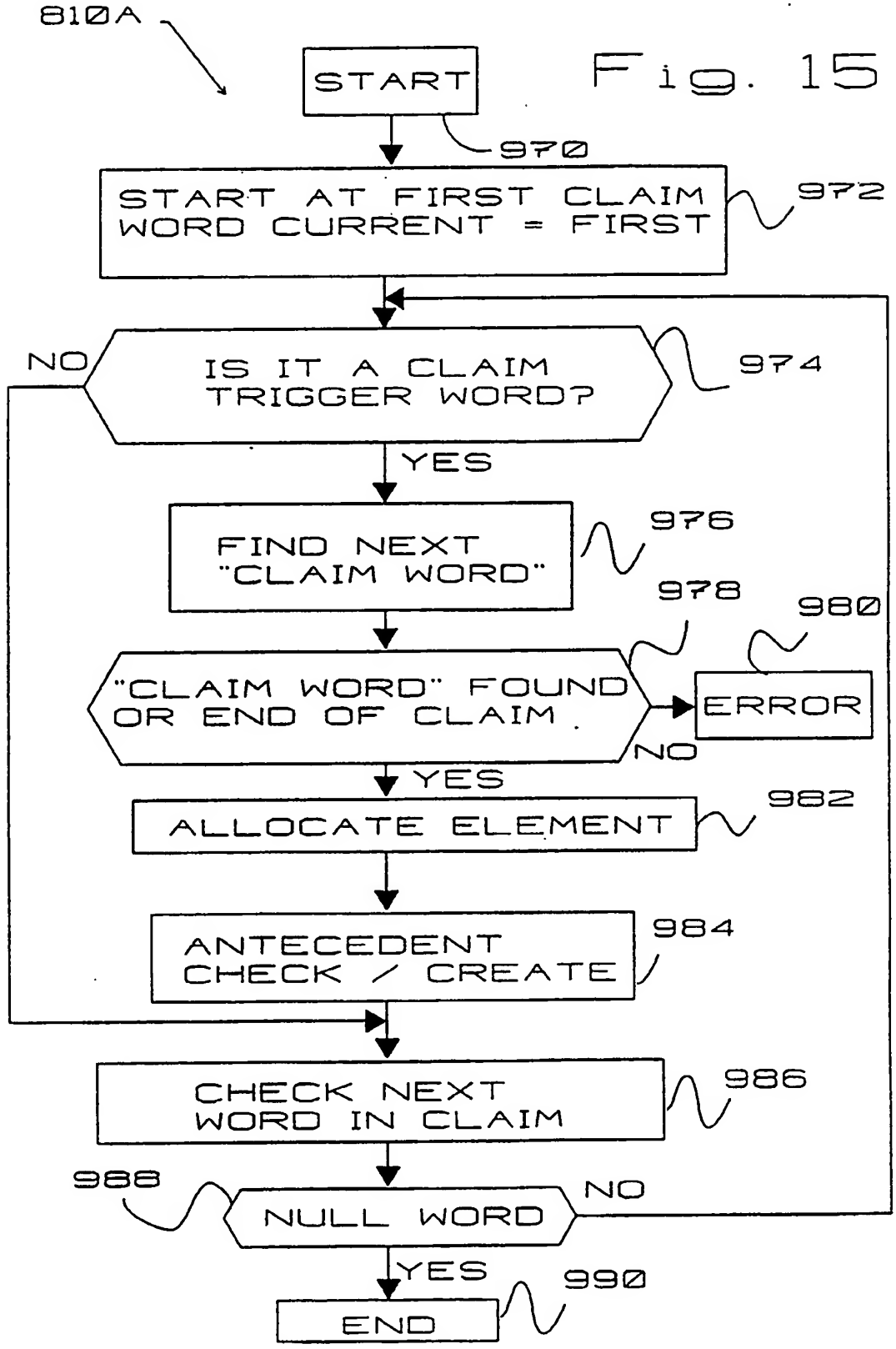

810A
START
970
START AT FIRST CLAIM WORD CURRENT = FIRST
972
NO
IS IT A CLAIM TRIGGER WORD?
974
YES
FIND NEXT "CLAIM WORD"
976
978
980
"CLAIM WORD" FOUND OR END OF CLAIM
ERROR
NO
YES
ALLOCATE ELEMENT
982
ANTECEDENT CHECK / CREATE
984
CHECK NEXT WORD IN CLAIM
986
988
NULL WORD
NO
YES
END
990

Fig. 15

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/05561

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :G06F 15/38
US CL :364/419.13

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 364/419.13; 382/7; 395/146

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	AUTOMATED PATENT SYSTEM (APS), MAY 1991 US DEPARTMENT OF COMMERCE, US PATENT AND TRADEMARK OFFICE, PAGES i, iii AND 27-32.	1-63
Y	MANUAL OF PATENT EXAMINING PROCEDURE, ORIGINAL FIFTH EDITION, AUGUST 1983, LATEST REVISION OCTOBER 1989, PAGES 600-29 TO 600-46 AND 700-7 TO 700-24.	1-63
Y	U.S.A., 4,773,009 (KUCERA ET AL.) 20 SEPTEMBER 1988. SEE ABSTRACT AND COL. 1 TO COL. 2, LINE 35.	1-63

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 30 JULY 1993	Date of mailing of the international search report OCT 06 1993
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer ROY ENVALL Telephone No. (703) 305-9711

Form PCT/ISA/210 (second sheet)(July 1992)*

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS (AUTOMATED PATENT SYSTEM)
S L1 (DOCUMENT# OR TEXT# OR FILE# OR WORD#) AND (RETRIEV? OR SEARCH?)
S L2 CHARACTER# AND SENTENCE# AND SECTION#
S L3 L1 AND L2
S L2 AND HEAD?
S L4 AND PUNCTUATION#